0

UNIVERSITY OF GENOVA

PHD PROGRAM IN BIOENGINEERING AND ROBOTICS

Laser Scanning Microscopy with SPAD Array Detector: Towards a New Class of Fluorescence Microscopy Techniques

by

Giorgio Tortarolo

Thesis submitted for the degree of *Doctor of Philosophy* (32° cycle)

March 2020

Dr. Giuseppe Vicidomini — Supervisor

Prof. Giorgio Cannata — Head of the PhD program

Thesis Jury:

Prof. Ilaria Testa, *KTH Royal Institute of Technology* — External examiner

Prof. Jörg Enderlein, *Georg August University* — External examiner

Abstract

Laser scanning microscopy is one of the most common architectures in fluorescence microscopy. In a nutshell: the objective lens focuses the laser beam(s) and generates an effective excitation spot which is scanned on the sample; for each pixel, the fluorescent image is projected into a single-element detector, which – typically – spatially and temporally integrates the fluorescent light along its sensitive area and the pixel dwell-time, thus providing a single-intensity value per pixel. Notably, the integration performed by the single-element detector hinders any additional information potentially encoded in the dynamic and image of the fluorescent spot. To address this limitation, we recently upgraded the detection unit of a laser scanning microscope, replacing the single-element detector with a novel SPAD (single photon avalanche diode) array detector. We have shown at first that the additional spatial information allows to overcome the trade-off between resolution and signal-to-noise ratio proper of confocal microscopy: indeed, this architecture represents the natural implementation of image scanning microscopy (ISM). We then exploited the single-photon-timing ability of the SPAD array detector elements to combine ISM with fluorescence lifetime imaging: the results show higher resolution and better lifetime accuracy with respect to the confocal counterpart. Moreover, we explored the combination of our ISM platform with stimulated emission depletion (STED) microscopy, to mitigate the non-negligible chance of photo-damaging a sample. Lastly, we showed how the SPAD array-based microscope can be used in the context of single-molecule/particle tracking (SMT or SPT) and spectroscopy. Indeed, we implemented a real-time, feedback based SMT architecture which can potentially correlate the dynamics of a bio-molecule with its structural changes and micro-environment, taking advantage of the time-resolved spectroscopy ability of the novel detector.
We believe that this novel laser scanning microscopy architecture has everything in its favour to substitute current single-element detector approaches; it will enable for a new class of fluorescence microscopy techniques capable of investigating complex living biological samples with unprecedented spatial and temporal characteristics and augmented information content.

Table of contents

List of figures

Nomenclature

Acronyms / Abbreviations

AOM Acusto-optic modulator

AOTF Acusto-optic tunable filter

APD Avalanche photo diode

APR Adaptive pixel reassignment

ASIC Application specific integrated circuit

BCD Bipolar-CMOS-DMOS

CCD Charge-coupled device

CLSM Confocal laser scanning microscopy

CMOS Complementary metal–oxide semiconductor

CoM Center of mass

CW Continuous wave

DCR Dark count rate

DFD Digital frequency domain

FCS Fluorescence correlation spectroscopy

FFT Fast Fourier transform

FIFO First in first out

FLIM Fluorescence lifetime imaging

FLISM Fluorescence lifetime image scanning microscopy

FPGA Field-programmable gate array

FRC Fourier ring correlation

FRET Förster resonance energy transfer

FWHM Full width at half maximum

GaAsP Gallium arsenide phosphide

GUI Graphical user interface

H2T Host to target

HyD Hybrid detector

ISM Image scanning microscopy

MLA Micro-lenses array

PCB Printed circuits board

PDE Photon detection efficiency

PMF Polarisation maintaining fiber

PMT Photo multiplier tubes

PR Pixel reassignment

PSF Point spread function

pSTED Pulsed STED

RTC Real-time component

SBR Signal to background ratio

SML Single molecule localisation

SMT Single molecule tracking

SNR Signal to noise ratio

SPAD Single photon avalanche diode

SPLIT Separation of photons by lifetime tuning

SPT Single particle tracking

STED Stimulated emission depletion

T2H Target to host

TCSPC Time correlated single photon counting card

TDC Time to digital converter

tE-PSF Temporal effective PSF

TTL Transistor–transistor logic

Chapter 1

Introduction

1.1 Fluorescence laser scanning microscopy

Far-field fluorescence microscopy is one of the most popular tools in the life sciences (Cella Zanacchi et al. (2014); Lichtman and Conchello (2005)). Its success roots on its rather exclusive advantages: by using (visible or near infra-red) light as the probe, it allows to investigate the sample under physiological conditions, feature of paramount importance to produce sound biological conclusions out of the experimental observations. Not less importantly, it grants specific and highly sensitive detection of the molecules of interest, by leveraging the fluorescence principle coupled with the opportune sample preparation. It is indeed possible to conveniently modify the sample - e.g., by immunostainining or by the expression of fluorescent fusion proteins (Griesbeck (2004); Zhang et al. (2002)) - to ensure that only the target molecules, within the usually vast pool of molecules constituting the sample, will be fluorescent. In this condition, after shining opportune light on the sample, only such designated molecules will contribute to the recorded fluorescence signal, resulting in fluorescence images with high specificity and excellent signal-to-background ratio (SBR). These two paramount advantages of fluorescence microscopy over other investigation techniques allowed to unravel many fundamental biological processes (Haustein and Schwille (2007)).

The broad family of far-field fluorescence microscopy techniques embraces two different approaches: wide-field (Verdaasdonk et al. (2014)) and laser scanning (or point scanning) microscopy techniques (Jonkman and Brown (2015)). Wide-field microscopes aim to shine the excitation light source on the whole field of view of interest, in order to excite and observe all fluorescence molecules simultaneously. Briefly, the excitation source is reflected by a dichroic mirror and focused on the sample thanks to an objective lens. It thus excites all the

fluorescent molecules within the field of view; the consequent fluorescence signal is then collected by the same objective lens, transmitted by the dichroic mirror to separate it from the excitation light, filtered by a detection filter and ultimately recorded by means of a pixellated detector, such as a camera, physically positioned in the optical path to image a magnified version of the sample, i.e. in the image plane of the optical system. Notably, all wide-field microscopy approaches intrinsically provide high frame-rate, given the parallel nature of the excitation strategy: all the field of view is illuminated and recorded at once. However, the extension of the illumination volume along the optical axis implies that each recorded image contains fluorescence information from a usually thick layer of the sample, hindering the chance of distinguishing features from different depths.

On the other hand, laser scanning microscopy techniques are based on a different platform. The core distinction relies in the excitation strategy: the desired field-of-view is no longer shone all at once, as for wide-field approaches, but it is scanned in a point-by-point fashion, by means of an opportune scanning system - e.g., a couple of galvanometric mirrors. For each desired point of the scan, the excitation light is focused by the objective lens to a (usually) diffraction limited volume in the sample, i.e., only to a tiny portion of the desired field-of-view. Consequently, only the molecules within this volume are excited and emit fluorescence, which is isolated thanks to a dichroic filter, and ultimately registered by a single element detector located in the image plane. The intensity information is stored, the excitation spot is moved to another point in the sample, and the fluorescence signal collected accordingly. The process is iterated until all the desired field-of-view is raster-scanned by the excitation spot; it is then possible to generate an image in which each pixel represents the fluorescence intensity value collected by the single element detector for each position of the excitation volume in the sample. Although the scanning approach leads intrinsically to longer acquisition times compared to wide-field architectures, the confinement of the excitation light to the diffraction limited spot paved the way for pivotal breakthroughs in the microscopy field, such as the development of Confocal Laser Scanning Microscopy (CLSM, Marvin Minsky (1961); Sheppard and Choudhury (1977), see Chapter 3), and Stimulated Emission Depletion microscopy (STED, Hell and Wichmann (1994), see Chapter 5).
In the context of this work, it is interesting to focus on one of the key components of any laser point scanning microscope: the photon detection hardware. The first implementations relied on photo multiplier tubes (PMTs), which are based on the photoelectric effect: briefly, the fluorescence photons strike the single element sensitive surface, a photocathode material; electrons are then ejected from the surface and then multiplied by the process of secondary emission. The amplitude of the output measured current is thus a function of the incident

photon flux. The PMT detectors are successfully used in many microscopy applications, since they show beneficial characteristics such as high gain, low noise, high dynamic range, and large area of collection. However, their usage severely limits the information content carried by the registered signal, in two distinct ways: (i) the analog nature of the amplification method implies the impossibility of straightforwardly measuring the time-of-arrival of incident photons: the output signal thus represents a temporal integration of the actual photon stream. Consequently, PMTs are not compatible with the vast family of assays leveraging the temporal information of detected photons, comprehensive of fluorescence lifetime imaging microscopy (FLIM, Becker (2012), see Chapter 4), separation of photons by lifetime tuning (SPLIT, Lanzanò et al. (2015), see Chapter 6), lifetime-based fluorescence resonance energy transfer (FRET, Förster (1965)), and others; (ii) moreover, the single element sensitive surface of PMT detectors *de facto* integrates the incident fluorescence photons over space, intrinsically hindering their spatial distribution. This information may be exploited to increase the spatial resolution and the signal-to-noise ratio (SNR) of the recorded images, as I will discuss in Chapter 3.
Choosing an alternative detector for the laser point scanning microscope allows to address one of these limitations at a time. In particular:

- The space-averaging can be avoided by the usage of complementary metal–oxide semiconductor (CMOS) or charge-coupled device (CCD) cameras: being pixellated detectors, cameras allow to sample the fluorescence signal with usually thousands of pixels. However, most currently available cameras are characterised by frame rates in the kHz regime, whilst the pixel-dwell time of a scanning microscope is in the MHz range – or below, thus, imaging speed is severely limited. Moreover, measuring the time-of-arrival of the fluorescence photons is impeded by the temporal integration within each pixel.

- On the other hand, single-photon detectors - such as avalanche photo diodes (APDs), single-photon avalanche diodes (SPADs) and hybrid detectors (HyDs) - allow to measure the actual number of incident photons: each detection event is indicated with a digital signal. Moreover, the temporal response of such detectors is usually extremely fast, granting a temporal resolution of the time-of-arrival of each photon in the order of hundreds of picoseconds. However, all single-photon detectors rely on a single element sensitive area - as PMTs - and thus hinder the spatial distribution of the fluorescence signal in the image plane.

At the time of this work, in the context of laser scanning microscopy, no detector allows to address both the temporal and the spatial integration limitations. Such a detector would allow to explore all the time-resolved spectroscopy-based approaches, allowing to greatly enrich the information content, including techniques able to investigate molecular dynamics, such as fluorescence correlation spectroscopy (FCS) and single-molecule tracking. Furthermore, it would also grant several benefits to the recorded images - mainly regarding the spatial resolution and the SNR - by leveraging the increased spatial information.

1.2 Aim of the work

This PhD thesis is focused on the introduction of a laser point scanning microscopy platform based on a novel, custom made detector: the single-photon-avalanche diode (SPAD) array detector (see Chapter 2). This detector stands as the synergistic combination of the two detector families described above, solving the temporal and spatial integration problem of laser point scanning microscope. As SPAD detectors, the SPAD array doesn't integrate the fluorescence photon stream over time: it indeed allows to count and to measure the time-of-arrival of each detected photon. On the other hand, similarly to cameras, the SPAD array presents a matrix of sensitive elements, thus sampling the fluorescence photons in space (Fig. 1.1).
This work aims to explore the various benefits granted by the novel laser point scanning microscopy platform, based on the SPAD array as the detection element. In particular:

- I investigate the resolution and SNR advantages obtained by using a pixellated detector in the context of point scanning systems, and I present our platform as the most straightforward implementation of Image Scanning Microscopy (ISM, Bertero et al. (1984); Müller and Enderlein (2010); Sheppard (1988), see Chapter 3);
- I leverage both the photon time-of-arrival and the spatial distribution information provided by the SPAD array introducing fluorescence lifetime image scanning microscopy (FLISM, see Chapter 4), to increase the resolution and the lifetime accuracy with respect to traditional FLIM imaging;
- I explore the advantages of our microscopy platform combined with STED microscopy (STED-ISM and pSTED-SPLIT, Chapter 5 and 6);

- I introduce a novel, real-time, feedback-based single molecule tracking (SMT) architecture (Chapter 7), which can potentially analyse the dynamics of the target bio-molecule leveraging the time-resolved spectroscopy ability of the SPAD array.

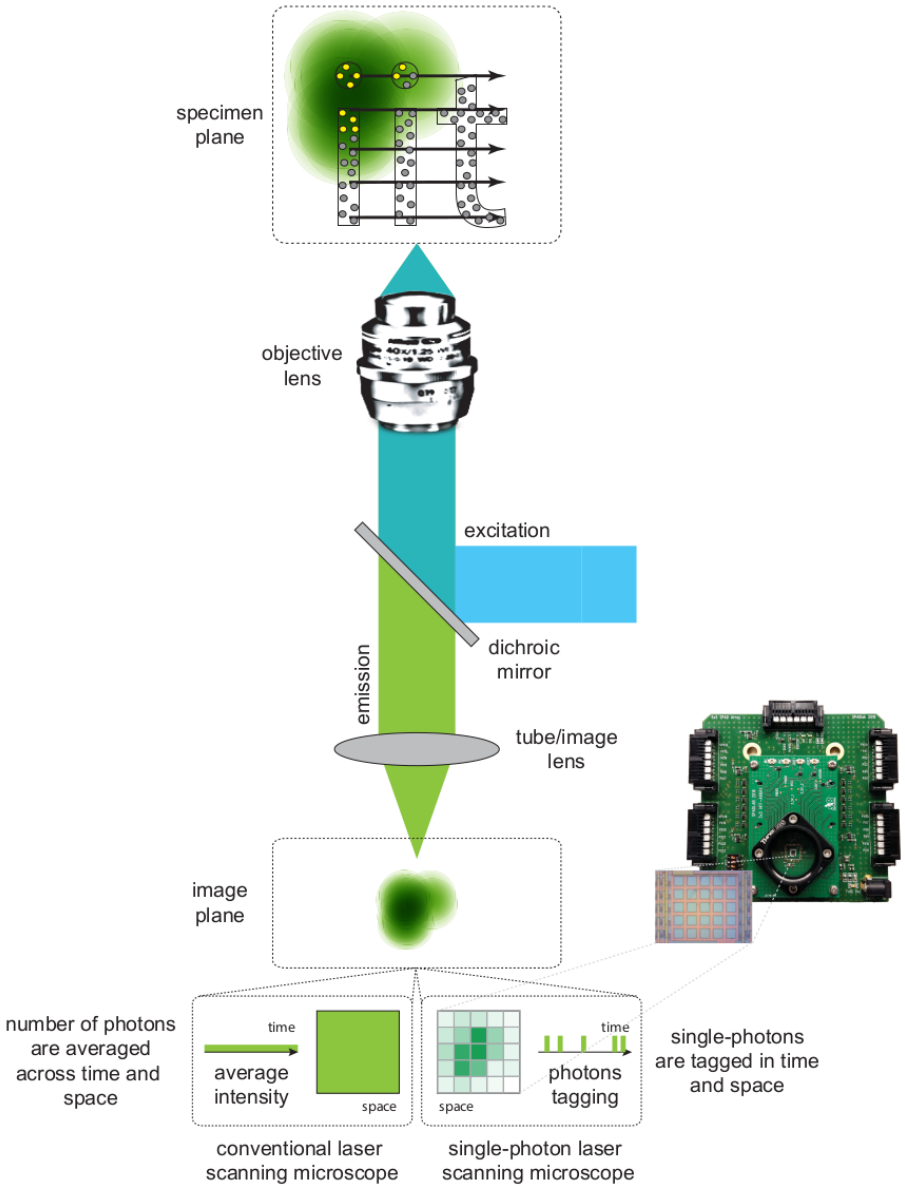

Figure 1.1 **Cartoon of the laser point scanning microscopy platform based on the SPAD array**. The excitation source is reflected by the dichroic mirror and focused by the objective lens on the sample, where it excites the fluorescent molecules. The resulting fluorescence light is collimated by the objective, transmitted by the dichroic mirror and ultimately focused by the tube lens to the image plane. The SPAD array, conversely to detectors for conventional laser scanning microscopes, allows to tag each photon in time and space.

Chapter 2

The SPAD Array Detector

2.1 Technical specifications

In the past years, single-photon avalanche diodes (SPADs) gained increasing popularity in the field of time-resolved spectroscopy, thanks to their photon-counting nature (Becker (2015); Zappa et al. (2007)). SPAD detectors are characterised by comparable - or even superior - performances with respect of photo-multiplier-tubes (PMTs), in terms of efficiency, noise and temporal response, and they additionally benefit from the great scalability, robustness, flexibility and reliability offered by the microelectronic fabrication technology. Indeed, the usage of multi-pixel microelectronic single-photon detectors contributes to reduce cost, overall size and system complexity. These advantages currently make the SPAD array one of the best technologies to develop a real photon-counting, two-dimensional detector array for fluorescence microscopy.

In this framework, we have designed and developed a novel single-photon detector array, specifically tailored for fluorescence microscopy applications Castello et al. (2019). The array is composed of a square matrix of 5×5 SPADs (Zappa et al. (2007)), having 75 μm distance (pixel pitch) and 50 μm side length (pixel size), with 5 μm corner radius (rounded-square active-area shape, as shown in Fig. 2.1a). This geometry results in a fill-factor (i.e., the ratio between photosensitive area and total detector area) of $\sim 44\%$, considering the external frame, otherwise $\sim 50\%$.

The overall detection system is composed of the custom-made SPAD array and two stacked printed circuits boards (PCBs), dedicated to the system management and the connectivity (Fig. 2.1a). The SPAD array application specific integrated circuit (ASIC) is fabricated using a 0.35 μm high-voltage CMOS technology by Fraunhofer IMS (Duisburg) and integrates the 25 SPAD detectors, which design is extensively described in Villa et al. (2014), together with

25 dedicated quenching and readout circuits, described in Bronzi et al. (2013). The ASIC provides 25 low-time-jitter digital outputs, whose rising edges are synchronous to photon detections. The first printed circuit board (called front-end board) hosts the detector ASIC - directly glued and electrically wire-bonded to it (circuit-on-board mounting technique) - together with programmable supply voltage generators (about 6 V for SPAD excess-bias, 3.3 V for readout and quenching circuits, about 25 V for SPAD biasing) and a digital-to-analog converter (DAC) for globally setting the hold-off time duration. The 25 digital output signals are passed to the second PCB (called connection board) through high-bandwidth connectors and then are fed to a set of 25 low-time-jitter buffers, able to drive 50 Ω impedance cables. These output cables are finally connected to an external FPGA-based system for collecting the photon counts. The connection board also includes a 8-bit microcontroller for system initialization and power management. Power supply comes from a single 5 V - 1 A external source. Individual pixels can be selectively enabled/disabled by the user using three external digital signals: two of them for serial communication (implementing a serial peripheral interface - SPI - communication) and one for the global enable. The detection system can be directly mounted on a multi-axis positional stage for a precise and reliable optical alignment. The choice of a relatively small number of pixels (i.e., 25) is based by previous theoretical studies, showing that a higher number of elements would provide only marginal improvements on the spatial resolution - for a fixed 1 A.U. projected-size of the detector array (Castello et al. (2015)). Notably, a higher number of elements would require more complex data-readout architectures, *de facto* hindering a fully parallel and independent reading of each pixel, and thus reducing both speed and versatility of the device. It is important to underline that each of the 25 elements of our SPAD array can deliver a fully-independent digital signal each time a photon is detected. As I will describe in the following Chapters, the asynchronous nature of both the SPAD array device and the FPGA-based data acquisition card allows for the most versatile data management architecture, and paves the way to the successful combination of the detection hardware with a number of fluorescence time-resolved spectroscopy/microscopy/tracking techniques.

Concerning the fill-factor, it is important to highlight that: (i) regarding all the imaging applications I will discuss in the following Chapters, the fluorescent photons are not uniformly spread across the whole detector, but indeed higher in the center and lower toward the periphery of the SPAD array - in case of a properly aligned system: thus, in the framework of imaging techniques, the odd number of elements of the detector makes the overall probability that a photon reaches the active area higher than the fill-factor itself; (ii) increasing the fill-factor, by reducing the spacing between elements, would likely deteriorate performances

in terms of optical cross-talk between adjacent pixels; (iii) the collection efficiency can be substantially improved by using a micro-lenses array (MLA) in front of the detector. We are currently working on the fabrication of a high fill-factor MLA directly on the SPAD array chip (Surdo et al. (2018)), anticipating an increase of the equivalent fill-factor to above 78% (i.e., above the theoretical value predicted by using a rectangular array of circular micro-lenses).

The SPAD array has been characterised in terms of photon detection efficiency (PDE), dark-count-rate (DCR), temporal response, optical cross-talk and after-pulsing probability. In Fig. 2.1b I show the measured PDE for the central pixel, as a function of wavelength and for different excess-bias voltages (V_{ex}) (other pixels exhibit similar performance). The PDE decreases when increasing the wavelength, ranging from about 45% at 480 nm down to 20-15% in the 600-700 nm region (both at 6 V_{ex} excess-bias). Higher PDE values could be achieved by using different fabrication technologies, as recently demonstrated with the Bipolar-CMOS-DMOS (BCD) technology (Sanzaro et al. (2018), see Appendix B). The dark count rate has been measured at 25 °C for each array element, resulting in an average DCR value around 200 counts *per* second (cps). The detector temporal response is shown in Fig. 2.1c for the central pixel only (similar results are obtained for all the 25 elements). It has been acquired using an external TCSPC board (SPC-630, Becker&Hickl) and a pulsed diode laser (32 ps of FWHM, 1 MHz of repetition rate and 850 nm of wavelength, Advanced Laser System), with all the other pixel turned ON, resulting in a time jitter below 200 ps (FWHM). The optical cross-talk probability between pixels is lower than 2% among closest neighbors (in the orthogonal direction). Finally, the afterpulsing probability ranges from 6.5% when enforcing a SPAD hold-off time of 50 ns, down to 1.4% with 200 ns hold-off time. Increasing the hold-off time is beneficial for the reduction of afterpulsing probability, but as a drawback, it correspondingly reduces the maximum count-rate of the detector. However, the maximum count rate of the SPAD array module is in general much higher than a single SPAD, thanks to the full parallel reading of each element and due to the fact that in a point-scanning microscope all fluorescent photons collected from a specific region of the sample are spread across all elements of the SPAD array. This property is also particularly important in the context of fluorescence lifetime, which historically has been strongly limited by the relative small maximum count-rate supported by single-photon timing detectors.

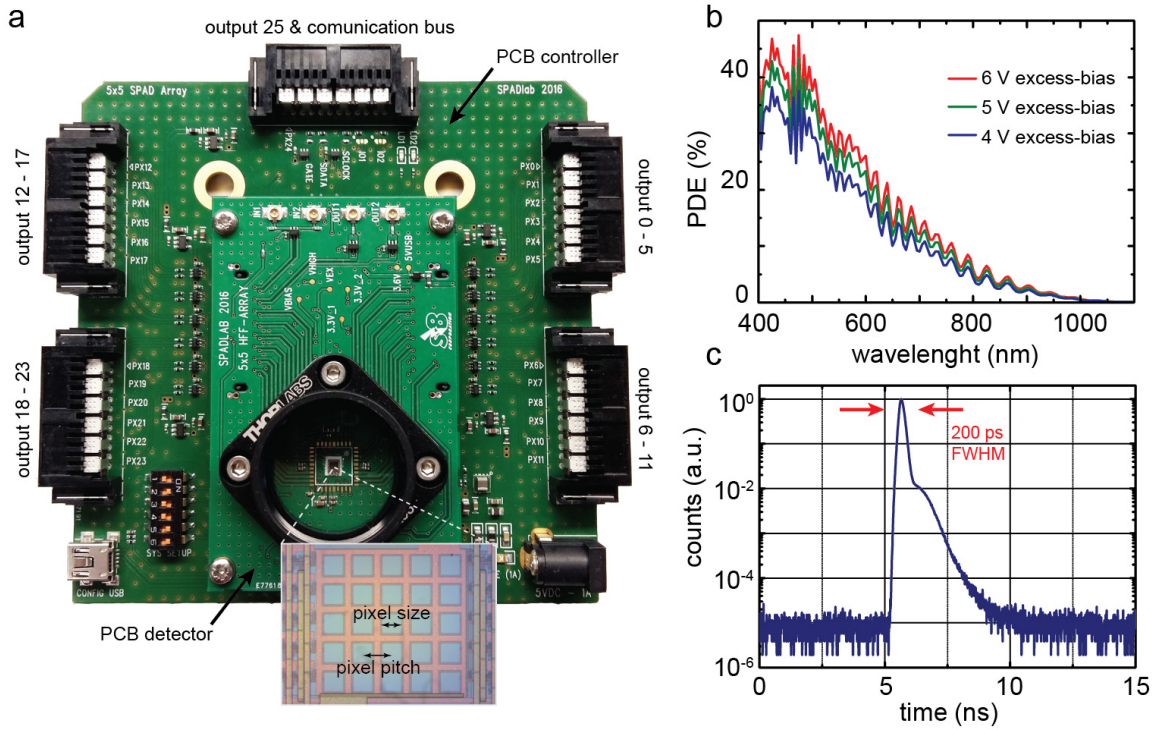

Figure 2.1 The SPAD array detector

2.2 Data acquisition and control system: "carma"

The choice of developing the SPAD array detector paved the way for the majority of projects that I will present in this Ph.D. thesis. It is important to underline here how its introduction led to a significant technological challenge, related mainly to the lack of an adequate solution to acquire the asynchronous signals generated by its 25 sensitive elements. In other words, at the time of the introduction of our SPAD array detector, to the best of our knowledge, no commercial or custom-made solution allowed to take advantage of its innovative features, namely the asynchronous acquisition of its independent signals. This challenge has been tackled and solved by developing the "carma" platform, which mainly manages the data acquisition and visualization from the SPAD array detector and controls the raster scanning of the microscope (Castello et al. (2019)). In this Section, I will briefly describe the core concepts of the "carma" platform, which I contributed to develop during the first period of my PhD work. For a full description of the "carma" platform see Castello et al. (2019).
The main features of the "carma" system can be divided into two general groups:

- control features: "carma" allows to control all the mechanical elements needed for the implementation of an optical scanning microscope: (i) the piezo scanning stages, for

the coarse positioning of the sample, and to move the objective lens along the optical axis to select the focal plane; (ii) the galvo-mirrors, to scan the laser beam(s) across the sample and to de-scan the fluorescent photons originated; (iii) the acusto-optic tunable filters (AOTFs), the acusto-optic modulators (AOMs) and the shutters, to block or to modulate the intensity of the laser beams;

- acquisition features: "carma" allows to acquire up to 26 digital signals in a fully parallel fashion and in synchronisation with the different actuators described above, thus enabling the interface with the SPAD array detector. Additionally, "carma" can acquire the analog tension signal typical of PMT detectors.

The architecture of the system is based on two main layers: carmaRTC (real-time component) and "carma" software. carmaRTC represents the hardware-layer of the system and performs real-time computing. carmaRTC runs on a field-programmable gate array (FPGA) board, the USB-7856R by National Instruments (NI), and it is written entirely using the LabView FPGA development environment from NI. The FPGA technology has been chosen mainly because of the complete control over the synchronization and timing of all signals and operations, and the possibility of custom on-board decision making that executes with hardware-timed speed and reliability. On the other hand, the high-level software is entirely written in C# using the IDE Visual Studio 2015. The software runs on the Microsoft .NET Framework (version 4.6). The software is responsible for the entire high-level job related to the microscopy task: from the user interaction (through a dedicated graphical user interface, GUI) to data storage, retrieval, management, image processing and all the computation required to create the image (for the part not in charge of the physical FPGA layer). Data can be saved in raw or Matlab format to perform image processing with external software.

Chapter 3

Image Scanning Microscopy with the SPAD Array Detector

3.1 Confocal laser scanning microscopy

In Chapter 1, I briefly introduced the family of laser point scanning microscopy techniques: as described, these approaches rely on a scanning architecture to span the excitation volume across the desired field of view; for each point in the scan, the fluorescence emitted by the fluorophores within the excitation volume is registered by a single element detector; an image is then build, in which any pixel reports an intensity value "proportional" to the fluorophore concentration in the associated sample volume.
In the context of laser point scanning microscopy, it is now interesting to discuss the spatial resolution, i.e., the minimum distance at which two separated objects are imaged as such: the resolving power of any given microscope is indeed one of its most prominent features, since it indicates the spatial scale at which details in the sample are distinguishable. For a scanning based architecture, the spatial resolution is strictly related to the extension of the excitation volume: intuitively, the smaller the excitation volume, the better the spatial resolution, and *vice-versa*. This relation is immediately clear considering that all fluorescent molecules lying within the excitation volume - and with similar optical properties - will be excited and will consequently emit fluorescence photons - with similar wavelengths, at any given point of the scanning. Hence, by observing only the fluorescence wavelength, it is impossible to distinguish their contributes to the fluorescence signal registered by the detector. To summarise, in the context of laser point scanning microscopy, and in absence of additional expedients, it is impossible to recognise as different two optically similar molecules that lie within the excitation volume.

In 1904, Ernst Abbe demonstrated that the dimension of the focal volume, and thus the spatial resolution of a point scanning microscope, is inferiorly limited by the diffraction of light (Abbe (1904)):

$$d_x = d_y = \frac{\lambda}{2n \sin \alpha} \tag{3.1}$$

where x and y are two orthogonal directions in the focal plane, λ is the wavelength of the excitation light, n is the refractive index of the medium and α is the maximum angle of the irradiated light captured by the objective. Defining the numerical aperture of the objective $NA = n \sin \alpha$, the size of the focal volume is mainly governed by the wavelength used and the numerical aperture of the optical system. Considering visible light and an objective with very high NA (λ = 650 nm, NA = 1.4), the dimension of the excitation spot - and thus the spatial resolution - is limited to $\sim$230 nm on the focal plane.

For decades, the diffraction of light has been considered an insurmountable limit for the spatial resolution of optical microscopes. However, the recent years have seen the development of the so-called super-resolution microscopy - or nanoscopy - techniques (Hell et al. (2015); Schermelleh et al. (2019)), which aim specifically to overcome the resolution limit imposed by diffraction, by different means. On one hand, Stimulated Emission Depletion Microscopy (STED, Hell and Wichmann (1994)) - and all the derived techniques - achieves super-resolution within the framework of a laser point scanning microscope by reducing the effective excitation spot - i.e., the volume from which molecules are allowed to emit fluorescence - by leveraging the stimulated emission process thanks to an additional light source. On the other hand, single molecule localisation (SML) techniques achieve similar results in the context of wide-field microscopy, by neglecting the simultaneous emission of molecules lying within the diffraction spot: the acquisition of enough temporal frames allows to localise each single emitter, and hence leads to the reconstruction of the super-resolution image.

These techniques were indeed the first to provide spatial resolutions well below the diffraction limit. However, the first microscopy technique to overcome the diffraction limit - although of a modest margin if compared to the properly called nanoscopy techniques - was confocal laser scanning microscopy (CLSM, Sheppard and Choudhury (1977)). The confocal microscope is based on the previously described laser point scanning optical architecture, although with the introduction of a physical pinhole in front of the detector. Positioning the pinhole in the optical path has two main consequences: (i) photons generated from out-of-focus planes are physically filtered by the pinhole, which avoid them to reach the sensitive area of the

detector. Therefore, the confocal microscope provides images with much improved signal to background ratio, in which only fluorophores from the desired focal plane are registered. This property is often referred to as the optical sectioning capability of the confocal microscope. (ii) In the condition of a point-like pinhole aperture, the spatial resolution of the confocal microscope is improved of a factor of $\sqrt{2}$ over the diffraction limit; in other words, the spatial cut-off frequency is doubled with respect to the conventional non-confocal laser scanning microscope.
The excellent optical sectioning capability of confocal microscopy greatly contributed to the popularity of the technique, now one of the most diffused in the field of the life sciences. On the other hand, the resolution improvement over the diffraction limit is rarely exploited, since it is theoretically achieved only if the pinhole is extremely closed. In this configuration, only few photons actually reach the detector, resulting in a heavy reduction of the signal to noise ratio (SNR) of the recorded image. As a consequence, users usually open the pinhole to achieve a satisfactory SNR, but doing so gradually decreases - and eventually nullifies - the super-resolution effect. In the following sub-chapters I will analytically describe such a trade-off between resolution and signal to noise ratio, in the context of the image formation process of a confocal microscope, and I will introduce a possible solution to overcome it, the Image Scanning Microscopy technique.

3.2 Image formation in a confocal laser scanning microscope

In this Section, I will analytically describe the image formation process of a confocal laser scanning microscope (CLSM), underlying the trade-off between resolution and signal to noise ratio. To this aim, it is important to observe that a CLSM can be described as a linear and spatial invariant system, and as such it can be completely characterised by its impulse response function. In the context of fluorescence microscopy, the impulse response function may be thought as the intensity distribution recorded by the imaging system when observing a point-like emitter, e.g., a fluorescent molecule. Intuitively, the light produced by such molecule cannot be recorded as a point again - because of diffraction - but it spreads out among a relative large area: the diffraction spot. The intensity distribution $k(x)$ of this spreading represents the impulse response function of the imaging system, called point-spread function (PSF) in the context of microscopy. For sake of simplicity, I will describe the PSF as a 1D function, with x a single lateral coordinate. The PSF consists of a pronounced central

main maximum surrounded by side maxima of much less intensity, the so-called Airy disks. Although in theory the PSF is described by Bessel functions, very often it is approximated by a Gaussian distribution (Harke et al. (2008)). As mentioned, the PSF completely describes the performances of an imaging system: notably, the PSF is related to the maximum achievable spatial resolution, often indicated as its full-width-at-half-maximum (FWHM) value or, in case of the Gaussian approximation, to its variance σ^2.

To analytically derive the PSF of a confocal system, it is convenient separate the illumination and the detection path. To describe the intensity distribution of a laser beam focused by the objective lens, it is possible to derive (Vicidomini (2008)) an integral representation for the intensity distribution generated by a lens transforming a monochromatic plane wave in a converging spherical wave. In the framework of laser scanning microscopy, this distribution intensity is generally called illumination point spread function h_{ill}, since represents the illumination spot used to excite a portion of the specimen. The same integral representation can be used to compute the image of a point source, since it can be modelled as a divergent spherical front wave collected by the lens and transformed into a plane wave. The image distribution intensity generated by a point source is generally called detection point spread function h_{det}, since it represents the image of the point in the detection plane of the system. Notably, it is possible to formulate the PSF of an imaging system simply in terms of illumination and detection PSFs. To this aim, we assume that the illumination and the detection are limited by finite sized apertures, a_{ill} and a_{det} respectively. In a confocal microscope, the illumination aperture is determined by the illumination pinhole, and the detection aperture is determined by the size of the detection pinhole. In general, a point in the object is illuminated through the illumination aperture a_{ill} of the light source. The light source can be modelled as a collection of point sources spread over the aperture, and with constant intensity C. Each point source is projected by the objective on the sample. The illumination intensity can therefore be written as the convolution between the illumination aperture and the diffraction point spread function:

$$i_{ill}(x) = (Ca_{ill}(x)) \star h_{ill}, \tag{3.2}$$

where $\star$ denotes the convolution product and h_{ill} the illumination PSF at the illumination wavelength λ_{ex}. Assuming that fading and saturation fluorescence effects can be neglected, the emitted light i_{em} is proportional to the illumination intensity i_{ill}, to the fluorescence molecules concentration in the sample, represented by the function $f(x)$, and to the brightness

B of the fluorescence molecules:

$$i_{em}(x) = i_{ill}(x) f(x) B. \tag{3.3}$$

The emitted light is imaged by the same objective onto a photo-sensitive sensor. Behind the objective, the detection aperture limits the emission light with respect to the sensor. The intensity behind this aperture can be modeled as:

$$i_{det}(x) = a_{det}(x)(h_{det}(x) \star i_{em}(x)). \tag{3.4}$$

The above formulations are valid for the general fluorescence microscope, and they can be tuned to different microscopy configurations by choosing the appropriate values for the illumination and detection apertures. In the case of confocal microscopy, a formulation for the signal i_{det} can be obtained assuming that both the excitation and the detection pinholes are infinitely small - condition which is formally modeled by using two Dirac functions $\delta(x)$. Combining the previous equations, we can therefore write the detection intensity i_{det} for a given point of the scan:

$$i_{det}(x') = \delta(x') \left(h_{det}(x') \star \left((h_{ill}(x') \star \delta(x')) f(x') \right) \right) \tag{3.5}$$

The detection intensity is recorded by a single element detector, resulting in a single intensity value of i_{SED} for each point of the scan. Such value can be obtained by integrating over the surface of the sensor - assumed as infinitely large:

$$\begin{aligned} i_{SED}(0) &= \int_{-\infty}^{+\infty} i_{det}(x') dx' = \\ &= \int_{-\infty}^{+\infty} \delta(x') \left(h_{det}(x') \star \left((h_{ill}(x') \star \delta(x')) f(x') \right) \right) dx' = \\ &= \int_{-\infty}^{+\infty} \delta(x') \left(h_{det}(x') \star \left((h_{ill}(x')) f(x') \right) \right) dx' = \\ &= \left(h_{det} \star (h_{ill} f(x')) \right)(0) = \int_{-\infty}^{+\infty} h_{det}(-\xi) h_{ill}(\xi) f(\xi) d\xi. \end{aligned} \tag{3.6}$$

Since the bi-dimensional image in a confocal laser scanning microscope is obtained by scanning the specimen, the intensity recorded by the photo-sensor for each point can be derived from Equation 3.6 by shifting both the illumination and detection over the object:

$$\begin{aligned} i_{SED}(x) &= \int_{-\infty}^{+\infty} \delta(x'-x)\Big(h_{det}(x') \star \big((h_{ill}(x') \star \delta(x'-x))f(x')\big)\Big)dx' = \\ &= \int_{-\infty}^{+\infty} \delta(x'-x)\Big(h_{det}(x') \star \big((h_{ill}(x'-x))f(x')\big)\Big)dx' = \\ &= \int_{-\infty}^{+\infty} \delta(x'-x)\Big(\int_{-\infty}^{+\infty} h_{det}(x'-\xi)h_{ill}(\xi-x)f(\xi)d\xi\Big)dx' = \\ &= \int_{-\infty}^{+\infty} h_{det}(x-\xi)h_{ill}(\xi-x)f(\xi)d\xi = \big(h_{det}(x)h_{ill}(-x)\big) \star f(x). \end{aligned} \tag{3.7}$$

Substituting the object function with the Dirac function leads to the PSF of a confocal laser scanning microscope:

$$h_{CLSM}(x) = h_{det}(x)h_{ill}(-x) \tag{3.8}$$

Notably, the previous formulation is valid only in the condition of an infinitely closed pinhole. To extend this result to real scenarios, we replace the Dirac function used to describe the pinhole in Eq. 3.7 with the finite size aperture a_{det}:

$$\begin{aligned} i_{SED}(x) &= \int_{-\infty}^{+\infty} a_{det}(x'-x)\Big(h_{det}(x') \star \big((h_{ill}(x') \star \delta(x'-x))f(x')\big)\Big)dx' = \\ &= \int_{-\infty}^{+\infty} a_{det}(x'-x)\Big(\int_{-\infty}^{+\infty} h_{det}(x'-\xi)h_{ill}(\xi-x)f(\xi)d\xi\Big)dx = \\ &= \int_{-\infty}^{+\infty} \Big(\int_{-\infty}^{+\infty} a_{det}(x-x')h_{det}(x'-\xi)dx'\Big)h_{ill}(\xi-x)f(\xi)d\xi = \\ &= \int_{-\infty}^{+\infty} (a_{det} \star h_{det})(x-\xi)h_{ill}(\xi-x)f(\xi)d\xi = \\ &= \big((a_{det}(x) \star h_{det}(x))h_{ill}(-x)\big) \star f(x). \end{aligned} \tag{3.9}$$

Similarly as before, by imposing $f(x) = \delta(x)$ we obtain the PSF of a confocal laser scanning microscope with the finite pinhole aperture $a_{det}(x)$:

$$h_{CLSM}(x) = \big(a_{det}(x) \star h_{det}(x)\big)h_{ill}(-x). \tag{3.10}$$

Considering an aberration-free system, the even symmetry of $h_{ill}(x)$ leads to the well-known following formulation:

$$h_{CLSM}(x) = h_{ill}(x)\big(h_{det}(x) \star a_{det}(x)\big). \tag{3.11}$$

This equation clarifies the trade-off between resolution and signal to noise ratio introduced before. Let's consider - in a perfectly aligned sytem - the case for which $\lambda_{ill} = \lambda_{det}$ (i.e., in the absence of Stokes shift), and also let's consider both the illumination PSF h_{ill} and the detection PSF h_{det} as (equal) Gaussian distributions. In this condition, Eq. 3.11 can be written as follows:

$$h_{CLSM}(x) = G(0,\sigma^2)\big(G(0,\sigma^2) \star a_{det}(x)\big), \tag{3.12}$$

where $G(\mu,\sigma^2)$ indicates the Gaussian distribution centered in μ and with variance σ^2. If now we consider a confocal system with an infinitely opened pinhole ($a_{det}(x) = 1, \forall x$), we see that its PSF is only governed by the diffraction limited illumination PSF, here $G(0,\sigma^2)$. This system would hence provide images with standard resolution, but of high signal to noise ratio granted by the infinitely opened pinhole. Conversely, if we consider the condition for which the pinhole is infinitely closed ($a_{det}(x) = \delta(x)$), the PSF now equals $G^2(0,\sigma^2) = G(0,\frac{\sigma^2}{\sqrt{2}})$. In other words, the resolution of such a system is improved over the diffraction limit of a factor $\sqrt{2}$, at the cost of a theoretically null signal to noise ratio.

3.3 Overcoming the trade-off between resolution and SNR

In 1988, Colin J.R. Sheppard hinted at the possibility of further increasing the resolution of a confocal microscope by replacing the single element detector with a detector array (Sheppard (1988)). This idea was suggested by his previous works, in which he noted that displacing the detector in a confocal system was unexpectedly providing an image of improved resolution, but of a point displaced from the optical axis. He wrote: "[...] This suggests a method of further improving the resolution. The signals from all points of a detector array are measured, but instead of integrating directly as in conventional imaging, each signal is reassigned to its particular image point".

In other words, Sheppard proposed to substitute the single element detector - traditionally used in conventional confocal microscopy - with a bi-dimensional array of detectors. He envisaged that in this condition each element of the detector array would capture an image with improved resolution, of a slightly misplaced point: indeed, the sensitive area of each element of the detector array acts - under the correct optical configuration - as a very closed

virtual pinhole, and thus grants the resolution improvement. The significant difference of such a system when compared to a traditional confocal microscope relies in the photons that are "rejected" by the virtual or physical pinhole: in a confocal microscope, they are actually lost, resulting in a decreased signal to noise ratio of the recorded image; in the system proposed by Sheppard, conversely, they are collected by other elements of the bi-dimensional detector array. The result - after a single scan of the excitation light on the sample - is a collection of super-resolved images, one per element of the detector array; each of them represents shifted views of the same sample, given the displacement of the corresponding sensitive element with respect from the optical axis. Notably, integrating the signal from all the sensitive elements, for each pixel of the scan - i.e., summing all the acquired images - leads to the traditional confocal result, in the configuration for which the pinhole aperture is as big as the whole sensitive area of the detector array: hence, the resolution improvement is lost. Notably, Sheppard suggested that it would be otherwise possible to shift each image of the correct amount, and only then performing the sum. This operation, known as "pixel reassignment", would preserve the improved resolution of each of the single images, while increasing dramatically the SNR. In this framework, the system proposed by Sheppard stands as an improvement over the traditional confocal microscope, since it potentially overcomes the trade-off between resolution and SNR.
To better analyse the image formation process in such a system, we may describe the PSF of the i_{th} element of detector array by leveraging the previously introduced equation 3.11, expressing the single *PSF* of an aberration-free confocal system:

$$h_i(x) = h_{ill}(x)\Big(h_{det}(x) \star a_{det}(x)\Big), \tag{3.13}$$

where x is the lateral spatial coordinate, $h_{ill}(x)$ is the illumination PSF, $h_{det}(x)$ describes the detection PSF, $\star$ indicates the convolution operation and $a_{det}(x)$ describes the virtual pinhole aperture, i.e., the sensitive area of the i_{th} element of the detector array. For the sake of this analysis, and similarly to what done in the previous sub-chapter, let's consider the following approximations:

1. both $h_{ill}(x)$ and $h_{det}(x)$ are represented as normalised Gaussian distributions, centered in the optical axis: $h_{ill}(x) = G(0, \sigma^2_{ill})$ and $h_{det}(x) = G(0, \sigma^2_{det})$;

2. the Stokes shift is null, i.e., the excitation and emission wavelengths are equal: $G(0, \sigma^2_{ill}) = G(0, \sigma^2_{det}) = G(0, \sigma^2)$;

3. the sensitive area of each element of the detector array, and thus the corresponding virtual pinhole aperture, is point-like: $a_{det}(x) = \delta(x - x_i), \forall i$. x_i represents the center of the sensitive area of the i_{th} element.

In these conditions, the convolution operation simply results in $G(-x_i, \sigma^2)$, and the equation reads:

$$PSF_i(x) = G(0, \sigma^2)G(-x_i, \sigma^2) \propto G(-\frac{x_i}{2}, \frac{\sigma^2}{\sqrt{2}}) \tag{3.14}$$

The result is hence a gaussian function, centered in $x_i/2$, with variance narrower by a factor of $\sqrt{2}$. Summing all the images corresponds to an equivalent confocal system for which the point spread function is $h_{sum} = \sum h_i \propto G(0, \sigma^2)$, leading to the loss of the resolution improvement. On the other hand, as introduced before, the pixel reassignment operation preserves it while enhancing the signal to noise ratio, by shifting back each recorded image of the correct amount. The system for which the final image is achieved by the operation of pixel reassignment is hence characterised by the following equivalent PSF:

$$h_{ps}(x) = \sum (h_i(x + \frac{x_i}{2})) \propto G(0, \frac{\sigma^2}{\sqrt{2}}) \tag{3.15}$$

To conclude, Sheppard proposed an improvement over the traditional confocal microscope, by substituting the single element detector with a detector array, and by processing the consequent pool of registered images *via* the pixel reassignment operation. This approach would lead to an increase of the spatial resolution and the signal to noise ratio with respect of confocal microscopy, and represents a possible solution for the related trade-off.

3.4 Image scanning microscopy: state of the art

As described in the previous sub-chapter, in 1988 Sheppard suggested a way to increase the spatial resolution and the SNR of a traditional confocal microscope, by leveraging the presence of multiple detectors and the pixel reassignment operation. However, the work stayed long overlooked, and no actual implementation was proposed, mostly because of technical limitations of the time regarding the acquisition hardware.
It was only twenty years later that the group of Jörg Enderlein proposed the first true implementation of this idea (Müller and Enderlein (2010)), which they named Image Scanning Microscopy (ISM). The name refers to the fact that, for each scanning point, an image is generated and recorded. This implementation, which relied on a conventional camera as

the detection unit, was successful in enhancing the resolution and the SNR with respect to confocal microscopy; however, the imaging speed of their architecture was limited by the low camera frame rate.
In the later years, various groups worked in the direction of mitigating this limitation: such efforts resulted in the introduction of multi-spot-excitation ISM architectures (Schulz et al. (2013); York et al. (2013)) and later by several all-optical ISM implementations (De Luca et al. (2013); Gregor et al. (2017); Roth et al. (2013); Winter et al. (2014); York et al. (2013)). In the context of all-optical ISM implementations, the final ISM image is obtained in real time, performing an additional scan on the detection arm and taking advantage of the integration time of the detection camera. These ingenious implementations have proven successful in retrieving the high frequency information usually hindered by noise in a confocal microscope, and in increasing the achievable imaging speed; moreover, the usage of CCD cameras is obviously convenient given the wide accessibility of the technology. However, the nature of the detection hardware is intrinsically hindering the temporal information of the time of arrival of each photon, since the photons are integrated in time within each frame. Hence, the potential combination of these ISM implementations with techniques leveraging the time-resolved spectroscopy information - e.g. fluorescence lifetime imaging - is not straightforward. Another plausible inconvenience is related to the optical image reconstruction: the success of this operation depends on the prior and precise knowledge of the shifts to be applied optically, here related to the ratio of the span of the two sets of scanners. Such shift values may be calculated *via* an opportune and laborious calibration, or derived from theory. However, it is worth mentioning here that the theoretical shift factor of $1/2$ - introduced in the previous sub-chapter - is only valid in the approximations of null Stokes shift and Gaussian excitation and emission *PSFs*. Any deviation from these conditions - which alters the shift factor - is in general difficult to model properly: on one hand, the spectral behaviours of different molecules strongly depend on the experimental environmental condition; on the other hand, the excitation and the emission PSFs are both sensitive to any misalignment of the optical system and to eventual system- or sample-induced aberrations. In other words, modelling properly *a priori* the shift values to be applied in the context of an all-optical ISM implementation may be, at best, non trivial.
More recently, the speed limit was solved in the AiryScan implementation of CLSM (Huff (2015)), in which an hexagonal bundle of optical fibers is connected to a one-dimensional array of GaAsP (gallium arsenide phosphide) photomultiplier tubes to mimic a detector array. However, this device does not synergistically integrate the temporal and spatial characteris-

tics of single-point detectors and cameras, which would be fundamental for versatile ISM implementations.

3.5 ISM with the SPAD array detector

In the previous Section, I have briefly presented various implementations of the ISM idea, which proved successful in overcoming the trade-off between resolution and signal-to-noise ratio. However, as discussed, all proposed approaches hinder the straightforward combination of ISM with time-resolved spectroscopy essays, and often require substantial modifications of the confocal optical setup.
Within this framework, we developed the SPAD array detector described in Chapter 2 to achieve what we believe to be the most straightforward implementation of ISM, offering various advantages with respect of previous works. (i) Firstly, our platform does not require significant modifications of the optical system. Indeed, the simple operation of substituting the single element detector of a confocal laser scanning microscope with our SPAD array - coupled with a simple lens to ensure the correct total magnification - upgrades the system to an image scanning microscope (Fig. 3.1a). (ii) The SPAD array allows to record the time-of-arrival of each detected photon: hence, our ISM is straightforwardly compatible with - and eventually beneficial to - time-resolved spectroscopy techniques such as fluorescence lifetime imaging (see Chapter 4). (iii) The asynchronous nature of both the SPAD array detector and the acquisition FPGA architecture leads to a versatile microscopy platform: it is indeed possible to acquire simultaneously all the images generated from the SPAD array after a single scan, one image *per* sensitive element (Fig. 3.1b). Notably, in this condition the final ISM result is obtained *a-posteriori*, conversely to the real-time reconstruction granted by all-optical ISM implementations. However, our strategy allowed us to develop what we call "adaptive pixel reassignment" (APR): the process of retrieving the shift values directly from the raw scanned images (Fig. 3.1c). Notably, this approach takes intrinsically into account the potentially unknown Stokes shift of the observed molecule, any sample - or system - induced aberration, and eventual optical misalignments; for these reasons, in our experience, it always outperforms the traditional pixel reassignment, which conversely relies on theoretically modelled shift values. (iv) Our implementation of the ISM does not alter any key component of a typical laser scanning microscopy system, other than the detection arm; hence, it paves the way for the straightforward combination with other imaging techniques based on a scanning system, e.g., Stimulated Emission Depletion microscopy (STED, see Chapter 5, 6) and two-photon excitation microscopy (Koho et al. (2019b)).

In the following Sections I will present in detail our ISM platform based on the SPAD array, and I will show its superior performances with respect to traditional confocal microscopy, in the context of fixed and live cell imaging.

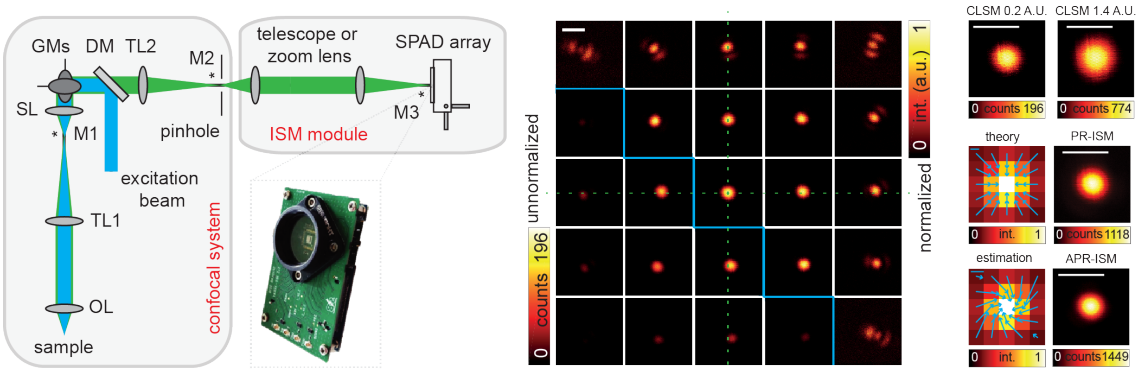

Figure 3.1 **The ISM implementation with the SPAD array**. (a) Schematic of the image scanning microscope. Excitation light (blue) passes the dichroic mirror (DM) and is deflected by the galvanometer mirrors (GMs). The pivot point of the scanner is projected by the scan and tube lenses (SL, TL1) into the back aperture of the objective lens (OL). Fluorescence (green) is collected by the objective lens, de-scanned by the GMs, filtered by the DM and projected by a second tube lens (TL2) into the pinhole plane. A telescope or a zoom lens is used to image the pinhole plane into the SPAD array and adds an extra magnification (from M2 to M3) to match the physical size of the detector to 1.4 A.U. The asterisks denote the planes conjugated. The picture shows the printed circuit board hosting the detector. (b) Matrix representing the scanned reflection images of a single isolated gold bead (80 nm). Each scanned image of the top-right corner and the bottom-left corner are normalized to its maximum intensity and to the maximum intensity of the central scanned image, respectively. Horizontal and vertical dashed-green lines guide the eye. Scale bar: 500 nm. (c) Side-by-side comparison of the PSF of the "ideal" confocal (0.2 A.U. pinhole), open confocal (1.4 A.U. pinhole), PR-ISM (i.e., pixel reassignment with the theoretical shift-vectors) and APR-ISM ISM (i.e., pixel reassignment with the estimated shift-vectors) obtained from (b). Scale bars: 500 nm. The fingerprint maps of (c) (i.e., the 5 x 5 image obtained integrating all the photons acquired during a scan by a single element of the detector array) superimposed with the theoretical and estimated shift-vectors projected in the image plane, are also shown. Scale bars: 100 nm.

3.6 Material and methods

3.6.1 Setup

We obtained the custom-based ISM system by modifying an existing point-scanning microscope. Briefly, the excitation beams were provided by two triggerable pulsed ($\sim$ 80 ps pulse-width) diode lasers (LDH-D-C-640 and LDH-D-C-485, Picoquant) emitting at 640 nm and 485 nm; we controlled their intensity/power by two acousto-optical modulators (AOM, MT80-A1-VIS, AA opto-electronic). After being combined with a long/short pass dichroic mirror (F43-491, AHF Analysentechnik) and reflected by a multi-band dichroic mirror (ZT 405-488-594-640-775, Chroma), the two excitation beams were deflected by two galvanometric scanning mirrors (6215HM40B, CT Cambridge Technology) and directed toward the objective lens (CFI Plan Apo VC 60×, 1.4 NA, Oil, Nikon) by the same set of scan and tube lenses used in a commercial scanning microscope (Confocal C2, Nikon). The fluorescence light was collected by the same objective lens, descanned, and passed through the multi-band dichroic mirror as well as through a suitable fluorescence band pass filter, depending on the imaging experiment (685/70 nm or 525/50 nm, AHF Analysentechnik). A 300 mm aspheric lens (Thorlabs) focuses the fluorescence light into the pinhole plane generating a conjugated image plane with a magnification of 300×. For ISM measurements, the pinhole is maintained completely open. A telescope system, built using two aspheric lenses of 100 mm and 150 mm focal length (Thorlabs), conjugates the SPAD array with the pinhole and provides an extra magnification factor. The final magnification on the SPAD array plane is 450×, thus the size of the SPAD array projected on the specimen is $\sim$ 1.4 A.U. (at the far-red emission wavelength, i.e. 650 nm) and $\sim$ 1.7 A.U. (at the green emission wavelength, i.e. 525 nm). Every photon detected by any of the 25 elements of the SPAD array generates a TTL signal that is delivered through the dedicated channel (one channel for each sensitive element of the detector) to an FPGA-based data-acquisition card (NI USB-7856R from National Instruments), which is controlled by our own data acquisition/visualization/processing software, *carma*. The software-package *carma* also controls the entire microscope devices needed during the image acquisition, such as the galvanometric mirrors, the axial piezo stage, and the acousto-optic modulators (AOMs), and visualises the data. In particular, carma synchronises the galvanometric mirrors with the photon detection to distribute photons between the different pixels of the images. All power values reported for this setup refer to the power measured at the sample plane.

3.6.2 Image reconstruction algorithms

Pixel reassignment

To reconstruct the high-resolution ISM image we implemented the pixel-reassignment (PR) method, which consists of (i) shifting each scanned image (i, j) of a shift-vector $\mathbf{s}_{i,j}$; (ii) adding up all the shifted images. The shift is implemented in the Fourier domain, also allowing for sub-pixel shift vectors $\mathbf{s}_{i,j}$. The shift-vectors can be estimated theoretically on the basis of the geometrical properties of the SPAD array and optical characteristic of the scanning microscope: theoretical shift-vectors represent the distance between the SPAD array elements and the "central" element, scaled by the magnification on the detector, the pixel size of the image and the PR factor, which is equal to 2 in the case of reflection (illumination and detection PSF are identical).
However, in practice the shift vectors calculated using this model are significantly different from the real ones (Fig. 3.2). As previously discussed, a first source of deviation is the Stokes shift, i.e. the excitation and emission PSFs are not identical, but they have different width, meaning that the position of the maximum of the effective PSF is not exactly located midway the excitation PSF and the shifted emission PSF, and the PR factor is different from 2. A PR factor compensating for the Stokes shift can be estimated *a-priori* (Roth et al. (2013)). Another important source of deviation is the different aberrations which effectively change the shape of both the excitation and detection PSFs, and which are difficult to estimate *a priori*. These aberrations influence the PR factor, and more generally the shift vectors. For these reasons, we implemented a method to estimate the shift vectors directly from the series of scanned images. Clearly, this method is an *a posteriori* approach − not a real-time approach, as in the all-optical ISM implementations, where the final image is built-up pixel-by-pixel as a conventional confocal microscope − but it offers the important ability to compensate for system- and sample-dependent distortions. Notably, in the case of an all-optical implementation based on fluorescent re-scanning, two different PR factors can be implemented along the two-axis, but these factors need to be known *a-priori*.
We estimated the shift vectors $\mathbf{s}_{i,j}$ for the PR using a phase correlation approach, which is typically used to estimate the drift between two images. Before describing the phase correlation approach, we need to introduce the discrete notation for the scanned images. Indeed, images are usually acquired on a regular 2-dimensional raster scanning grid. If we identify each pixel by its index $\mathbf{n} = (n_x, n_y)$, we can denote the $N_x \times N_y$ scanned image as $\mathbf{g}_{i,j}(\mathbf{n})$ with $n_x = 1, ..., N_x$ and $n_y = 1, ..., N_y$.
Phase correlation estimates the shift between two similar images relying on a frequency-

domain representation of the data, which in our implementation is obtained through fast Fourier transform (FFT). To calculate the phase correlation between the two different scanned image $\mathbf{g}_{i,j}$ and $\mathbf{g}_{3,3}$, we first define the so-called correlogram $\mathbf{r}_{i,j}$

$$\mathbf{r}_{i,j} = FFT^{-1}\left(\frac{FFT\left(\mathbf{g}_{i,j}\right)FFT(\mathbf{g}_{3,3})^{*}}{\left|FFT\left(\mathbf{g}_{i,j}\right)FFT(\mathbf{g}_{3,3})^{*}\right|}\right), \tag{3.16}$$

and subsequently find the maximum of the correlogram, whose position denotes the drift/shift between the two scanned images $\mathbf{g}_{i,j}$

$$(\mathbf{s}^{x}_{i,j}, \mathbf{s}^{y}_{i,j}) = \underset{(\mathbf{n}_x,\mathbf{n}_y)}{\arg\max}(\mathbf{r}_{i,j}(\mathbf{n})). \tag{3.17}$$

The position of the maximum can be obtained using a Gaussian-fitting algorithm or a centroid algorithm in order to obtain sub-pixel values.
Interestingly, the shift vectors can be used to estimate the magnification M of the microscope on the detector array. In absence of Stokes shift, such as in reflection microscopy, the shift vectors are equal to half of the displacement value, $\mathbf{s}_{i,j} = \mathbf{d}_{i,j}/2$. Since the displacement value depends on the projected physical distances of the element of the SPAD array, which are well known values, the magnification can be calculated as

$$M = \frac{4p}{2\sum_{k_{x/y}=\{-1,1\}}\left|\mathbf{s}_{3+k_x,3+k_y}\right|}, \tag{3.18}$$

where p is the pixel-pitch, i.e. 75 μm in our SPAD array. Only the shift vectors linked to the first-order neighbours of the central element are used, since their estimation is more robust, i.e., they are associated to higher SNR scanned images. We used reflection imaging of gold beads and this approach to calibrate the magnification for all our experiments.

Multi-image deconvolution

Multi-image deconvolution is another approach for recombining the scanned images into an high-resolution image. In comparison to PR, deconvolution needs higher computational effort and prior information, such as the PSFs of the scanned images, but can provide higher SNR and higher effective resolution (Bertero et al. (2009)). Here, we derive the multi-image deconvolution algorithm following a maximum-likelihood (statistical) approach (Bertero et al. (2009)) and using a discrete notation for the object function, the PSFs and the digital

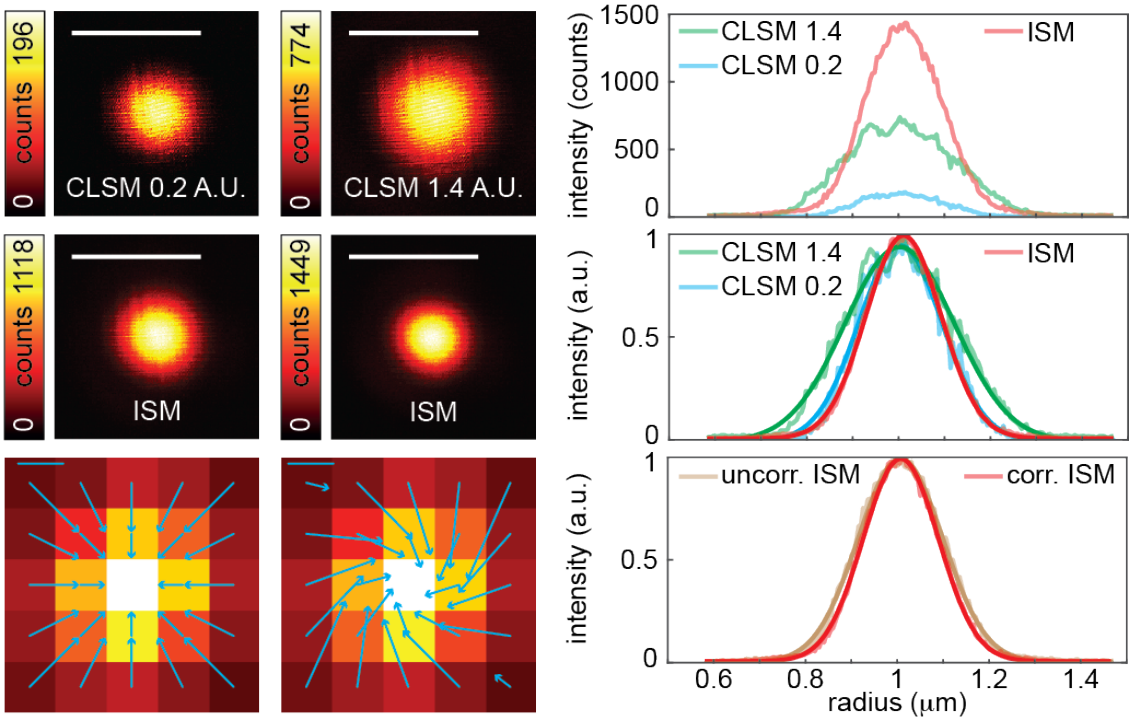

Figure 3.2 **Image scanning microscopy with adaptive pixel reassignment (APR)**. Side-by-side comparison of the effective PSFs for "ideal" confocal (0.2 A.U.), "open" confocal (1.4 A.U.), conventional PR-ISM (the pixel reassignment method uses theoretical shift vectors), and APR-ISM (the pixel reassignment method uses estimated shift vectors). The effective PSF is retrieved imaging in reflection modality a single isolated gold bead (80 nm). Scale bar: 500 nm. Bottom, fingerprint maps superimposed with the shift vectors for the conventional PR-ISM (left) and the APR-ISM (right). Scale bars: 50 nm. Right, radial PSFs obtained from their intensity profiles. The top panel shows the un-normalized PSFs, the middle and bottom panel show the normalized PSFs, together with the Gaussian fitted data. Pixel-dwell time: 50 μs. Pixel-size: 5 nm. Image format: 400 × 400 pixels. Scale bars: 500 nm.

images.

It is convenient here to describe the image formation process of the image scanning microscope: since a fluorescent image scanning microscope can be considered as a linear and space-invariant system, the relation between the expected (noise-free) scanned image $\bar{g}_{i,j}(\mathbf{x})$ – associated with the element (i, j) of the detector array – and the object/specimen function $f(\mathbf{x})$ can be described by a convolution operator $H_{i,j}$

$$\bar{g}_{i,j}(\mathbf{x}) = [H_{i,j}(f)](\mathbf{x}) = \int h_{i,j}(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y} = (h_{i,j} * f)(\mathbf{x}) \quad \text{with} \quad i = 1,..,5,\ j = 1,..,5, \tag{3.19}$$

where $h_{i,j}$ is the effective PSF associated to the element (i,j), $\mathbf{y}$ is the position in the sample and $\mathbf{x}$ is the position in the image back projected into the sample, i.e., the scanning position. Here, we consider a magnification equal to 1 between the object and image planes and a detector array with 5×5 elements. If we denote the discretized object function, the expected scanned images and the PSFs with the one-dimensional vectors $\mathbf{f}(n)$, $\bar{\mathbf{g}}_{i,j}(n)$, and $\mathbf{h}_{i,j}(n)$ (with $n = n_y N_y + n_x$), we can indeed write the image formation process as

$$\bar{\mathbf{g}}_{i,j} = \mathbf{H}_{i,j}\mathbf{f}, \tag{3.20}$$

where the $\mathbf{H}_{i,j}$ are the convolution matrices ($N_x N_y N_x N_y$ sized) associated with the convolution (linear) operator $H_{i,j}$ and the PSF $h_{i,j}$. Moreover, with the discretization of convolution integral of Equation 3.19 using cyclic convolution and periodic extension of the pixel values of $\mathbf{f}$ and $\mathbf{h}_{i,j}$ reduces $\mathbf{H}_{i,j}$ to a circular matrix, hence the transformation

$$\mathbf{H}_{i,j}\mathbf{f} = \mathbf{h}_{i,j} * \mathbf{f}, \tag{3.21}$$

can be easily computed by means of the FFT. Here and in all subsequent equations multiplication and division of one vector by another is taken to mean pixel-by-pixel.
The measurement process is dominated by shot noise and count rates are usually in the range of zero to a few hundred photons per pixel. Thus, for each pixel n and each scanned image (i,j), the measured value $\mathbf{g}_{i,j}(n)$ is the realization of a Poisson random variable with its expectation value given by $\bar{\mathbf{g}}_{i,j}(n)$. Because each pixel is statistically independent from the other, the probability of recording the series of scanned images $\mathbf{g}$ for a given specimen $\mathbf{f}$ is given by

$$P(\mathbf{g}|\mathbf{f}) = \prod_{(i,j)} \prod_{n} \mathrm{poi}[\mathbf{g}_{(i,j)}(n) | \bar{\mathbf{g}}_{(i,j)}(n)] = \prod_{(i,j)} \prod_{n} \frac{e^{-\left(\mathbf{H}_{i,j}\mathbf{f}\right)(n)} \left(\left(\mathbf{H}_{i,j}\mathbf{f}\right)(n)\right)^{\mathbf{g}_{(i,j)}(n)}}{\mathbf{g}_{i,j}(n)!}. \tag{3.22}$$

Since we assume knowledge of the probability density $P(\mathbf{g}|\mathbf{f})$ of the data and the specimen $\mathbf{f}$ appears as a set of unknown parameters, the problem of deconvolution can be approached as a classical problem of parameter estimation, which can be solved by the standard maximum

likelihood (ML) estimation approach. We introduce the likelihood function $\mathscr{L}_{\mathbf{g}}$, defined by

$$\mathscr{L}_{\mathbf{g}}(\mathbf{f}) = P(\mathbf{g}|\mathbf{f}), \tag{3.23}$$

which is a function of $\mathbf{f}$ only, since the series of scanned image $\mathbf{g}$ is given. Then, the ML-estimate of the unknown object $\mathbf{f}$ is any object $\mathbf{f}^*$ that maximize the likelihood function

$$\mathbf{f}^* = \arg\max_{\mathbf{f}} \mathscr{L}_{\mathbf{g}}(\mathbf{f}). \tag{3.24}$$

Since in our application the likelihood function is the product of a very large number of factors (Eq. 3.23 and 3.22), it is convenient to take the logarithm of this function; moreover, if we consider the negative logarithm, the maximization problem is transformed into one of minimization. By introducing the so-called discrepancy functional J, the deconvolution problem reads

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} -B\ln\mathscr{L}_{\mathbf{g}}(\mathbf{f}) + C = \arg\min_{\mathbf{f}} J(\mathbf{f};\mathbf{g}), \tag{3.25}$$

where B and C are suitable constants that can be introduced in order to simplify the expression of the functional. By using simple mathematics the discrepancy function of our algorithm reads

$$J(\mathbf{f};\mathbf{g}) = \sum_{i,j}\sum_{n}\left\{\mathbf{g}_{i,j}(n)\ln\frac{\mathbf{g}_{i,j}(n)}{\left(\mathbf{H}_{i,j}\mathbf{f}\right)(n)} + \left(\mathbf{H}_{i,j}\mathbf{f}\right)(n) - \mathbf{g}_{i,j}(n)\right\}. \tag{3.26}$$

For the solution of Equation 3.25 we chose the split-gradient-method (SGM, Vicidomini (2008); Vicidomini et al. (2009)) due to its robustness, the simplicity of its implementation and its capability to enforce non-negative constraint, i.e. $\mathbf{f} \geqslant 0$, in a natural fashion. For our discrepancy function (Eq. 3.26) the SGM iterations are given by

$$\mathbf{f}^{k+1} = \mathbf{f}^k \sum_{(i,j)}\left(\frac{\mathbf{1}}{\mathbf{H}_{i,j}^T\mathbf{1}}\mathbf{H}_{i,j}^T\frac{\mathbf{g}_{i,j}}{\mathbf{H}_{i,j}\mathbf{f}^k}\right), \tag{3.27}$$

where $\mathbf{H}_{i,j}^T$ is the transpose of the operator $\mathbf{H}_{i,j}$ and $\mathbf{1}$ is the vector whose entries are all equal to 1. In practice, the matrix-vector multiplication $\mathbf{H}_{i,j}^T\mathbf{1}$ generates a vector whose elements are the sum of $\mathbf{H}_{i,j}$ across its columns. Since the matrix $\mathbf{H}_{i,j}$ is cyclic, $\mathbf{H}_{i,j}^T\mathbf{1}$ is a vector whose

entries are all equal to the sum of the discretized PSF $\mathbf{h}_{i,j}$

$$w_{i,j} = \sum_n \mathbf{h}_{i,j}(n), \tag{3.28}$$

and the SGM algorithm (Eq. 3.27) reduces to

$$\mathbf{f}^{k+1} = \mathbf{f}^k \sum_{(i,j)} \left(w_{i,j}^{-1} \mathbf{h}_{i,j} \star \frac{\mathbf{g}_{i,j}}{\mathbf{h}_{i,j} * \mathbf{f}^k} \right), \tag{3.29}$$

where, for the sake of simplicity, we move to a vector notation and $\star$ denotes the correlation operation, which as with convolution can be implemented through a FFT. The algorithm in Equation 3.29 can be considered as an extension of the Richardson-Lucy (RL) algorithm for solving the multi-image deconvolution problem. Indeed, for a single image, the algorithm reduces to the well-known RL algorithm.
Finally, it is important to note that, throughout this thesis, we used a simplified Gaussian-based model to calculate the PSFs. For each element (i, j) we calculated a normalized (the integral is equal to 1) Gaussian PSF centered on $\mathbf{s}_{i,j}$ (Eq. 3.17). We used the same full-width at half-maximum for all the elements, but we scaled each PSF for a factor $w_{i,j}$ which takes into account the expected different SNRs of the associated scanned images. As scaling factors we used the values of the normalized fingerprint map. We estimated the FWHM directly from the images, by fitting the line intensity profile of single isolated sub-diffraction structure in the brighter scanned image with a Gaussian function. Notably, more advanced methods may be exploited to estimate the FWHM of each PSF: as an example, we recently leveraged the FRC analysis to characterise each PSF directly from the scanned images, hence reducing the *a priori* knowledge needed to successfully reconstruct the ISM result *via* image deconvolution (Koho et al. (2019a)).

3.7 Results

We have validated our Image Scanning Microscopy implementation based on the SPAD array by performing measurements of fluorescent calibration samples, fixed and live cells. Thanks to the versatility of our platform, we compared the ISM results of each experiment with the corresponding confocal images. Indeed, our implementation favoured this approach, since it allows to extract - from the very same measurement - a multitude of results:

- the image registered by the central element of the SPAD array. By considering only this contribute, we simulate a virtual confocal system which pinhole aperture equals the dimension of the single sensitive element ("ideal" or "closed pinhole" configuration). In the context of this work, and where not differently noted, the dimension of the corresponding virtual pinhole is 0.2 Airy Units;
- the sum of all the 25 registered images, without applying any shift. The result corresponds to simulating a virtual confocal system which pinhole aperture equals the whole SPAD array area ("open pinhole configuration", 1.4 Airy Units);
- lastly, the Image Scanning Microscopy results: the pixel reassignment ISM image (PR-ISM), the adaptive PR ISM image (APR-ISM), and the deconvolved ISM image (ISM^{++}), each obtained applying our reconstruction algorithms.

3.7.1 Fluorescent beads

We started evaluating the performances of our ISM platform by imaging fluorescent beads, a commonly used calibration sample. The results confirmed the expected superior performances of ISM with respect to confocal microscopy, for what concerns the spatial resolution and the signal to noise ratio. In particular, the higher spatial frequency information of the adaptive pixel reassignment and the deconvolved ISM imaging modalities resulted in clearly separated fluorescent beads, conversely to both the "ideal" and the "open pinhole" confocal configurations, as confirmed by line intensity profiling (Fig. 3.3).

3.7.2 Fixed cells

We proceeded to validate our ISM platform by performing various experiments on fixed cells. Firstly, we visually inspected the "ideal" confocal, "open pinhole" confocal, and APR-ISM images of tubulin filaments stained with Abberior STAR red: the resolution enhancement granted by the adaptive pixel reassignment approach resulted in clearly separated filament structures, conversely to both the confocal configurations (Fig. 3.4), as confirmed by line intensity profiling.
In order to better compare the effective spatial resolution content of the various imaging modalities, we leveraged the Fourier Ring Correlation analysis (FRC, see Appendix A), a technique that we previously extended to the case of point scanning laser microscopy (Tortarolo et al. (2018)). Very briefly, the FRC resolution metric is based on the concept of the effective cutoff frequency of a given microscopy system; conversely to other approaches

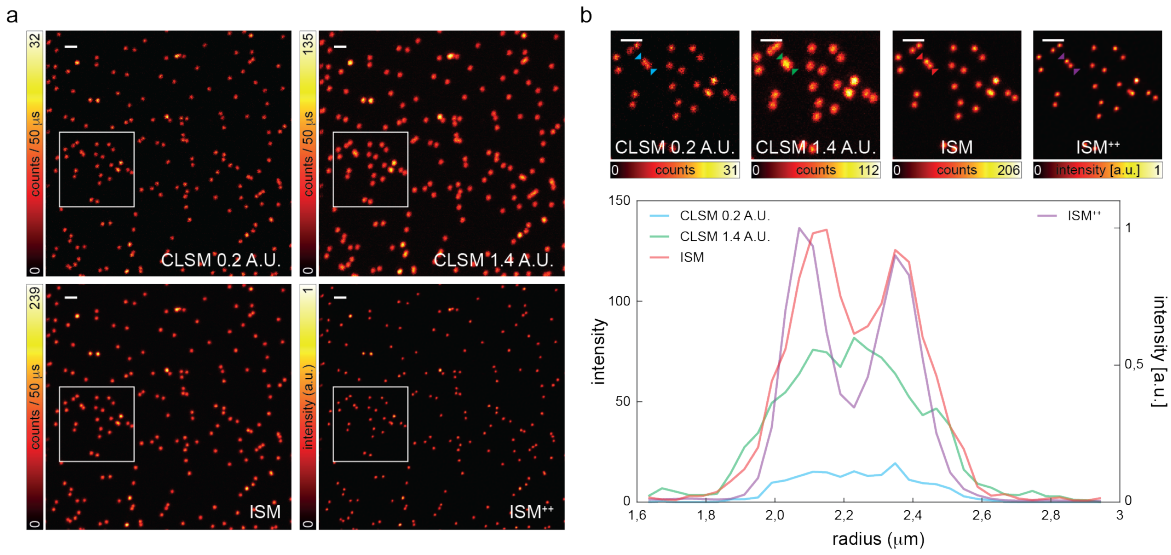

Figure 3.3 **Image scanning microscopy on fluorescent beads**. Image scanning microscopy on fluorescent beads. (a) Side- by-side comparison of "ideal" confocal, "open" confocal, APR-ISM, and deconvolved ISM++ (10 iterations) images (500 × 500 pixels, 40 nm pixel size) of 20nm red fluorescent beads. Pixel-dwell time: 50 μs. Pixel-size: 40 nm. Image format: 500 × 500 pixels. Excitation power P_{exc} = 56 nW. Scale bars: 1 μm. (b) Magnified views of the white boxes in (a). Scale bars: 1 μm. (c) Line intensity profiles across two close fluorescent beads at the position of the arrowheads in (b) for the different imaging modalities. Data are representative of n = 5 experiments.

(e.g., theoretical resolution estimation, calibration samples, full-width at half-maximum criterion) it is complete, straightforward, and sensitive to all factors that can affect the image resolution. Importantly, the resolution value retrieved by FRC analysis - referred to as "FRC resolution" - is highly dependent from the signal to noise ratio of the image. Hence, FRC analysis stands as an appropriate tool to compare ISM and confocal imaging modalities.
We performed FRC analysis on a series of Image Scanning Microscopy images of tubulin structures labelled with Atto647N in fixed HeLa cell, obtained increasing the excitation laser power (Fig. 3.5). FRC analysis confirmed the superior effective resolution of the ISM result when compared with both the "ideal" and "open pinhole" confocal configurations. Notably, this experiment shows the ability of APR-ISM to achieve the same resolution of “ideal” CLSM, but at only one tenth of the illumination intensity.
We additionally imaged the nuclear lamin of fixed HEK cell (Fig. 3.6), and large field-of-views of the nuclear scaffold structures, both stained with Alexa 488 (Fig. 3.7): in both cases, the APR-ISM approach led to images of increased resolution and signal to noise ratio, indicating the robustness of our method with respect to wide field-of-view imaging.

3.7.3 Live-cell image scanning microscopy

As discussed briefly at the beginning of this chapter, the non-invasive nature of fluorescence microscopy techniques greatly contributed to their wide success, allowing live-cell imaging to investigate dynamic cellular phenomena in their living environment. Confocal microscopy, in addition, grants images with a much increased signal to background ratio, thanks to its excellent optical sectioning capability. Moreover, opening or closing the pinhole allows the user to prioritise the signal to noise ratio or the spatial resolution, respectively: targeting resolutions below the diffraction limit requires a very closed pinhole, condition that potentially results in a severe decrease of the SNR of the recorded image. Increasing the power of the excitation laser source could mitigate this effect, allowing to gather enough fluorescence signal while conserving the super-resolution effect; however, the increased photon flux irradiating the sample may also lead to photo-bleaching of the fluorophores (and eventually to photo-toxicity effects), hindering or neglecting the chance of observing dynamic processes for extended amount of time.
By overcoming the trade-off between resolution and signal to noise ratio, Image Scanning Microscopy relaxes the live-cell imaging conditions, allowing sustained observation of live cell phenomena. To prove this, we performed long time-lapses of tubulin labelled live cells, which resulted in little to none photo-bleaching of the fluorophores. Notably, switching to

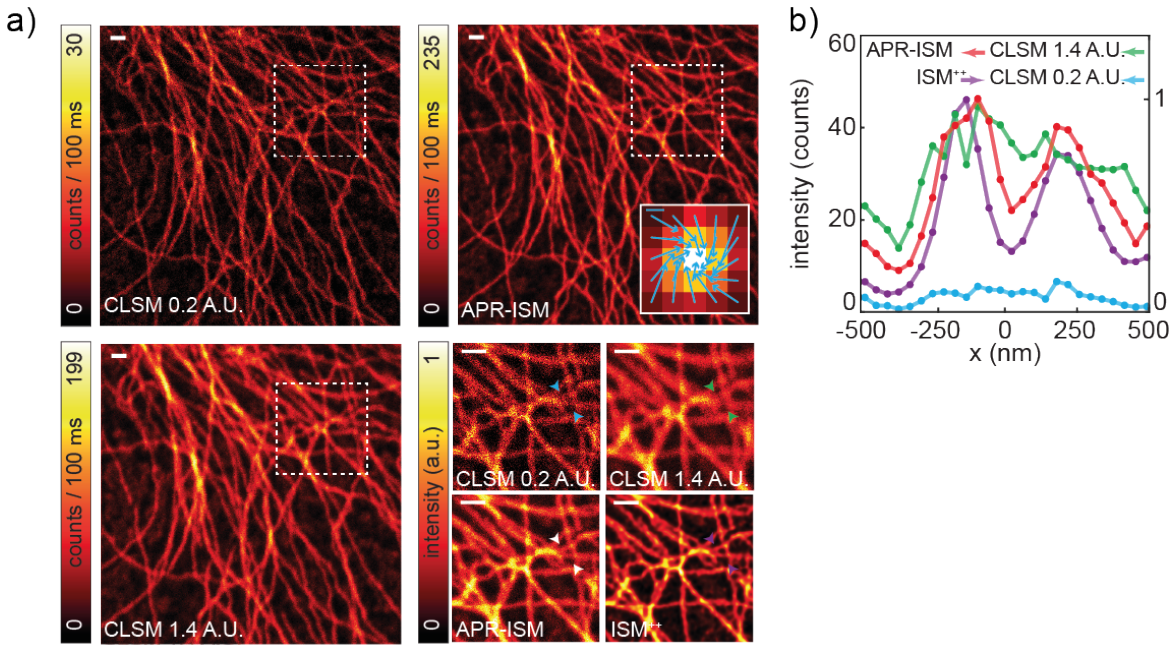

Figure 3.4 **ISM imaging of tubulin structures**. (a) Side-by-side comparison of ideal confocal, open confocal, and APR-ISM images of tubulin filaments stained with Abberior STAR red. Inset shows the fingerprint map and the estimated shift-vectors. Scale bar: 100 nm. Magnified views of the dashed box area are also reported, together with the ISM image obtained using multi-image deconvolution (ISM++, 5 iterations). Scale bars: 1 μm. Data are representative of n = 10 experiments. (b) Line intensity profiles across two branching tubular filaments at the position of the arrowheads for the different imaging modalities.

the equivalent "ideal" confocal modality, it was necessary to increase 10-fold the excitation laser power to achieve the same target resolution, in accordance to the previous experiments. This severe increase led rapidly to the photo-bleaching of the dyes, neglecting to perform a time-lapse of the same duration (Fig. 3.8). This experiment suggests that Image Scanning Microscopy is preferable to traditional confocal microscopy in the context of live-cell imaging.

To further test the compatibility of our ISM implementation with live cell imaging, we captured extended time lapses of tubulin filaments labelled in human HeLa cells (Fig. 3.8). We then observed mitochondria within a living cell, using the MitoTracker Deep Red label: the mobility of such organelles suggests the absence of any photo-toxic effect.

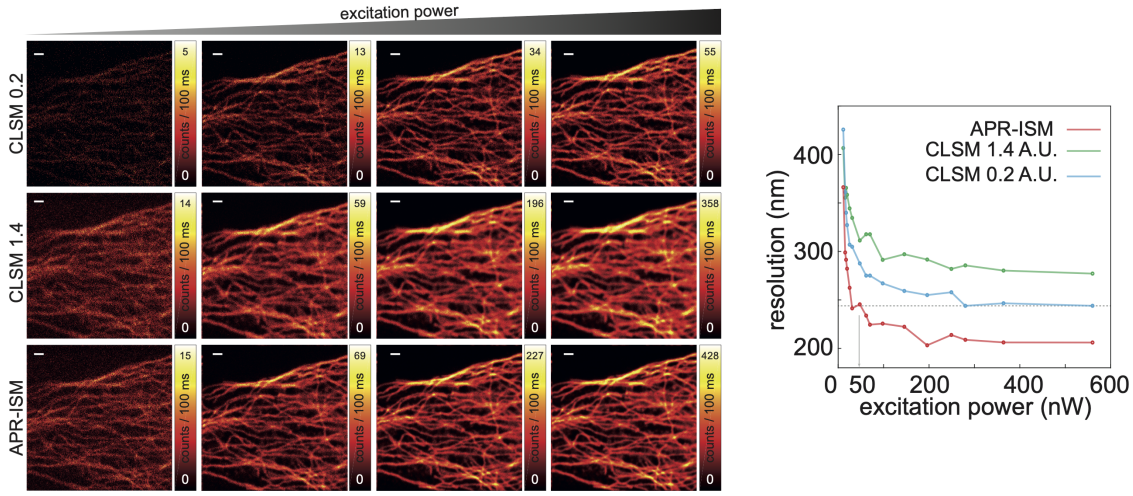

Figure 3.5 **FRC-based resolution scaling on ISM with increasing excitation power**. Series of tubulin (labeled with Abberior STAR red) images of the same area for "ideal" confocal (top), "open" confocal (middle), and APR-ISM (bottom) as a function of the excitation beam power. Bleaching was negligible across the whole imaging experiments. Pixel-dwell time: 100 μs. Pixel-size: 37,5 nm. Image format: 400 × 400 pixels. Excitation power Pexc = 50, 55, 62, 70, 90, 110, 170, 220, 250, 350, 520, 700, 890, 1000, 1300, 2000 nW. Scale bars: 1 μm. Right, Fourier ring correlation resolution as a function of the excitation power for the three different imaging modalities. Data are representative of n = 2 experiments.

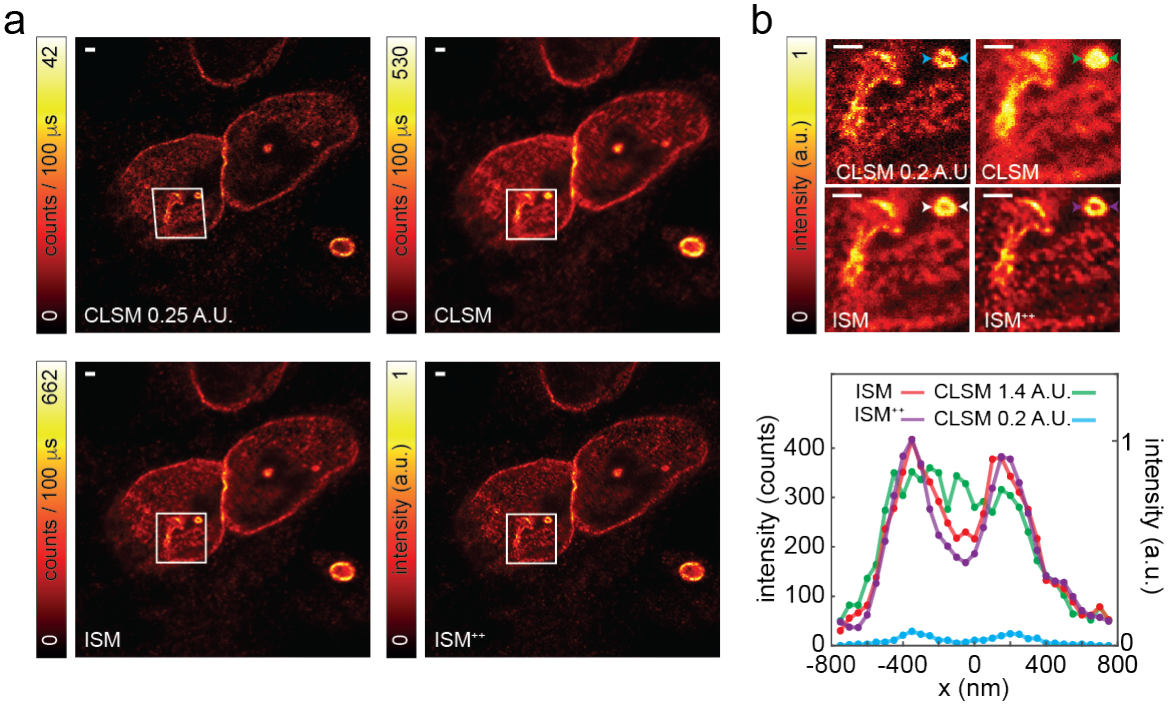

Figure 3.6 **Image scanning microscopy on nuclear lamin**. (a) Side-by- side comparison of HEK cell images of nuclear lamin stained with Alexa 488: "ideal" confocal (0.25 A.U), open confocal (1.7 A.U.), APR-ISM and deconvolved ISM++ (5 iterations), respectively. Image format: 600 × 600 pixels; (b) Magnified views of the white boxes in (a). (c) Line intensity profiles across a nuclear invagination (nucleoplasmic reticulum) at the position of the arrow-heads for the different imaging modalities. For all images: Pixel-dwell time, 100 μs; Pixel-size, 50 nm; Excitation power, Pexc = 840 nW. Scale bars: 1 μm. The FRC-based resolution values for the images are 238 nm, 263 nm, and 206 nm for "ideal" confocal, "open" confocal and APR-ISM, respectively. Data are representative of n = 10 experiments.

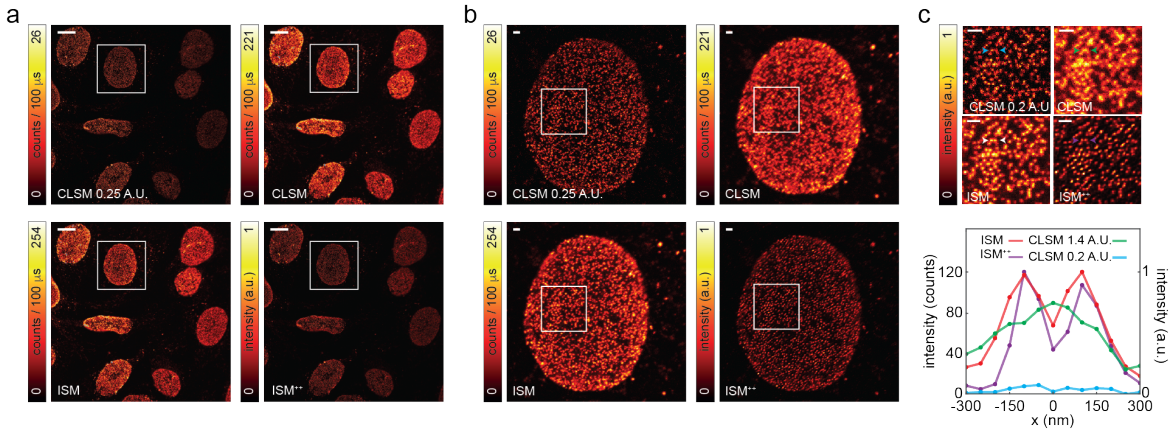

Figure 3.7 **Image scanning microscopy on nuclear pore complexes (NPCs).** (a) Comparison of large field-of-view images of the nuclear pore scaffold structures stained with Alexa 488: "ideal" confocal (top) and APR- ISM (bottom). Image format: 1500 × 1500 pixels; Scale bars: 10 μm. (b) Side-by-side comparison of single-cell images, "ideal" confocal (0.25 A.U), open confocal (1.7 A.U.), APR-ISM and deconvolved ISM++ (5 iterations), respectively. Notably, to highlight the property of the proposed ISM implementation to generate high-resolution large field-of-view image, all images are a digital zoom of the white boxes in (a). Image format: 400 × 400 pixels; Scale bars: 1 μm; (c) Magnified views of the white boxes in (b). Image format: 100 × 100 pixels; Scale bars: 1 μm. (d) Line intensity profiles across two closely packed NPCs at the position of the arrowheads for the different imaging modalities. For all images: Pixel-dwell time, 100 μs; Pixel-size, 66.6 nm; Excitation power, Pexc = 840 nW. The FRC-based resolution values for the 1500 × 1500 pixel images are: 213 nm, 253 nm and 207 nm, for "ideal" confocal, open confocal and APR-ISM, respectively. Data are representative of n = 10 experiments.

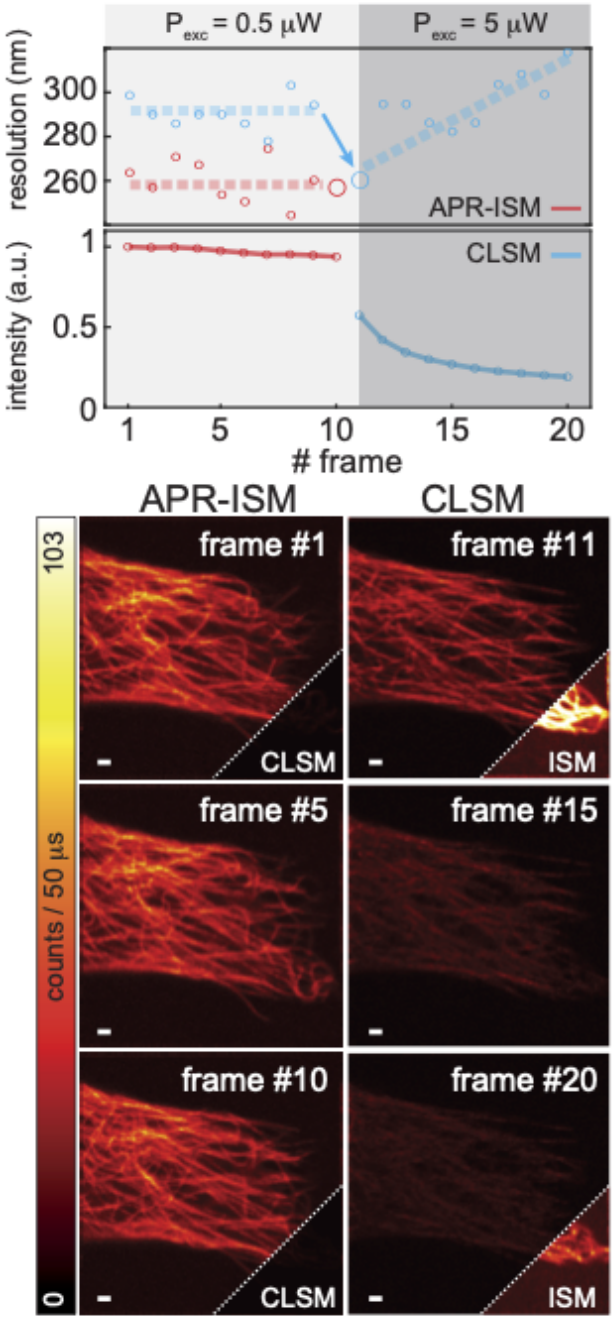

Figure 3.8 **Time-lapse Image Scanning Microscopy**. Time-lapse (4 minutes, 20 frames) imaging of SiR-tubulin in Human Hela live-cells (n = 4 experiments). Resolution and total fluorescence intensity (normalised at the first frame) as a function of the frame (top) for CLSM (1.4 A.U.) and APR-ISM. The first 10 frames are collected using 0.5 μW excitation beam power; at the 11^{th} frame the excitation beam power is increased to 5 μW in order to have similar resolution between CLSM and ISM (large circle). Representative frames of the time-lapse (bottom). Scale bars: 1 μm.

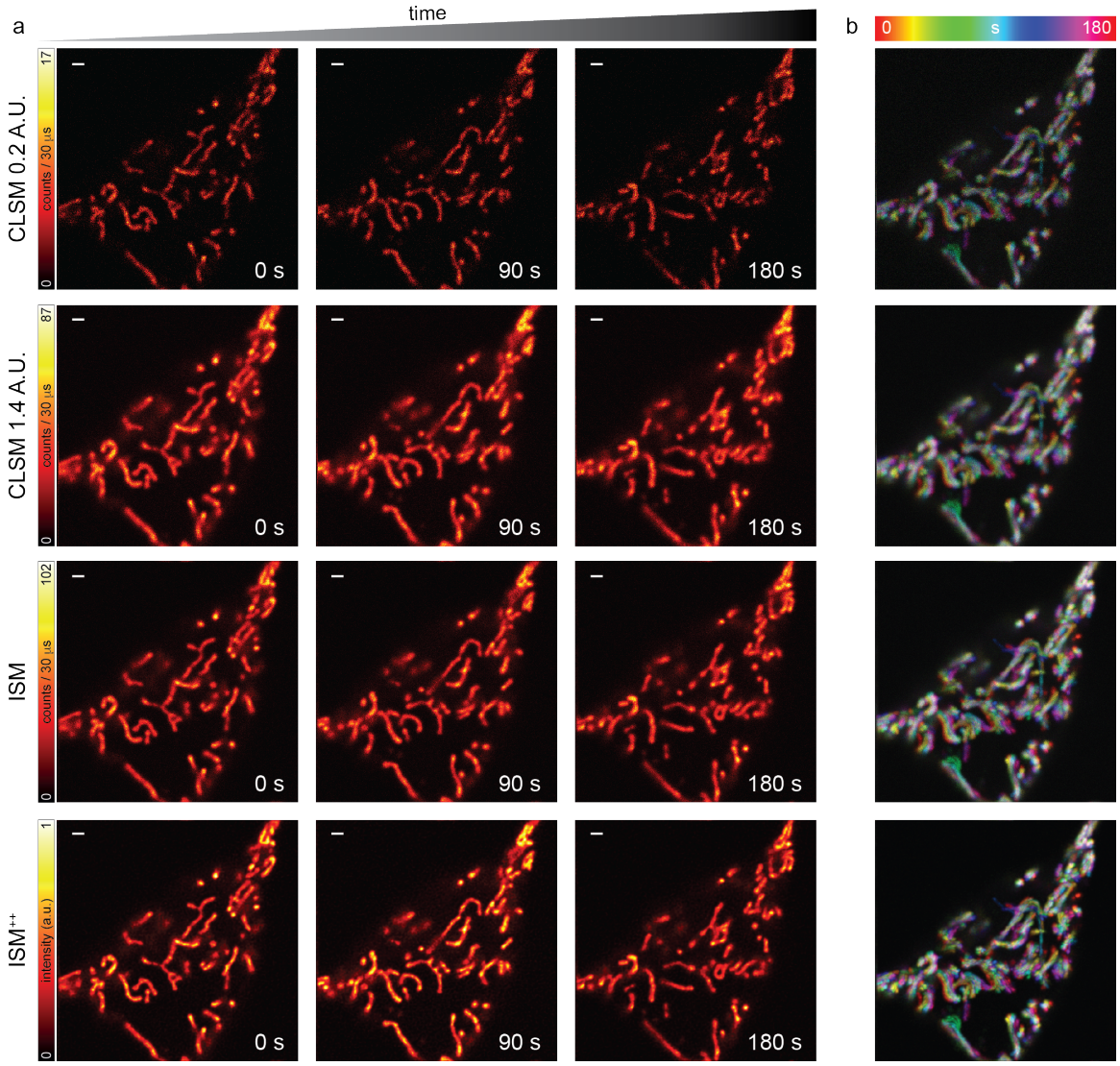

Figure 3.9 **Image scanning microscopy for live-cell mitochondria imaging**. (a) Side-by-side comparison of "ideal" confocal (top), "open" confocal (middle-top), APR-ISM (middle-bottom), and deconvolved ISM++ time-lapse (3 minutes, 24 frames) of mitochondria labelled with MitoTracker Deep Red in a living cell. For each imaging modality, three representative frames are shown (0 s, 90 s, and 180 s). Pixel-dwell time: 30 μs. Pixel-size: 40 nm. Image format: 500 × 500 pixels. Excitation power Pexc = 140 nW. Scale bars: 1 μm. (b) Maximum intensity projections (color-coded by time) of the time-lapses for the different imaging modalities. These projections allow to identify the fraction of mitochondria with minimal mobility (white) from the mobile one (colored). Data are representative of n = 10 experiments.

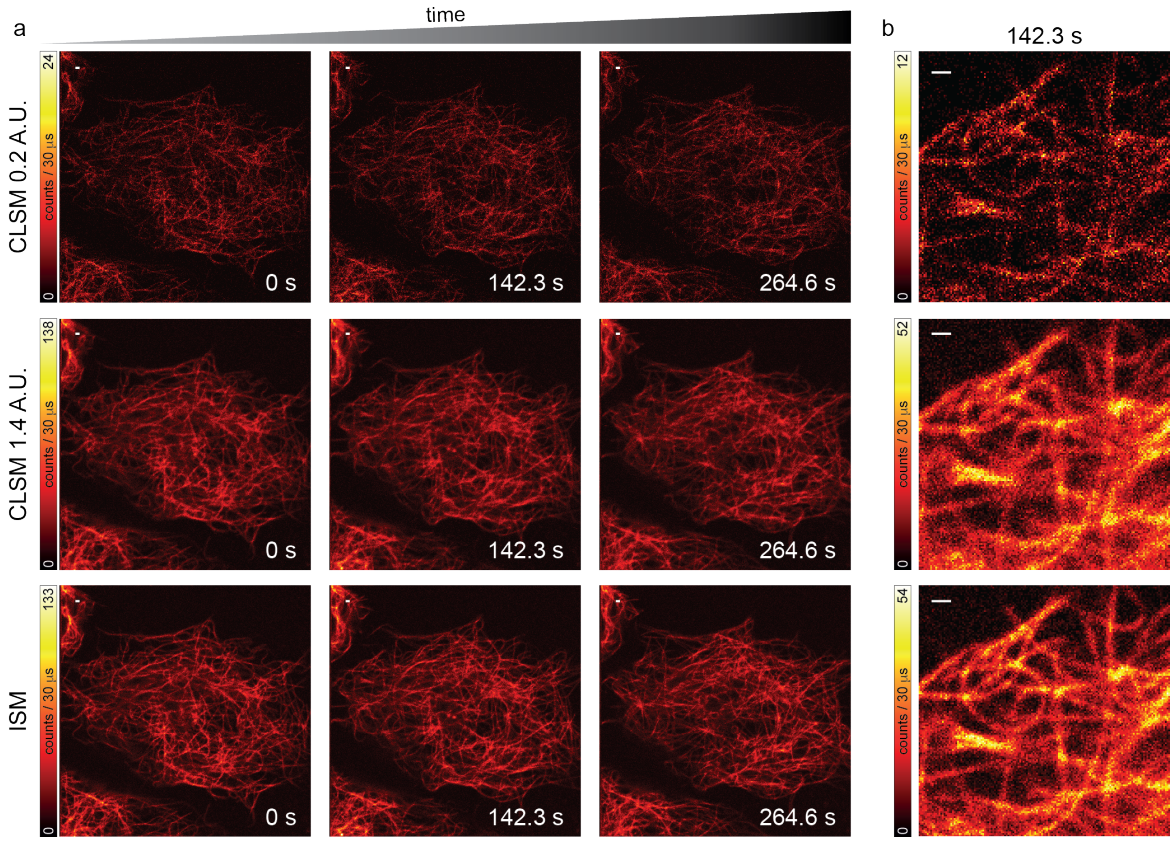

Figure 3.10 **Image scanning microscopy for live-cell tubulin imaging**. (a) Side-by-side comparison of "ideal" confocal (top), "open" confocal (middle), and APR-ISM (bottom) time-lapses (4 minutes and 30 seconds, 19 frames) of tubulin labelled with SiR Tubulin in a Hela live-cell. For each imaging modality, three representative frames are shown (0 s, 142 s, and 264 s). All frames are normalized in intensity to the maximum and minimum of the first frame of the respective time-lapse. This allows to appreciate the negligible photobleaching. Pixel-dwell time: 50 μs. Pixel-size: 57.2 nm. Image format: 700 × 700 pixels. Excitation power Pexc = 2 μW. Scale bars: 1 μm. (b) Magnified view of the white dashed boxes reported in (a). Data are representative of n = 3 experiments.

Chapter 4

Fluorescence Lifetime Image Scanning Microscopy

4.1 Fluorescence lifetime imaging microscopy

In Chapter 3, I described the image scanning microscopy technique, and our architecture based on the SPAD array detector. As discussed, the ISM result is a fluorescence intensity image: each pixel represent the fluorescence signal detected for the corresponding position of the excitation spot on the sample. Hence, ignoring saturation effects, the intensity of each pixel is proportional to the concentration of the targeted fluorescent molecule. Notably, as per all the fluorescence intensity imaging techniques, assessing quantitative information from ISM measurements may not be trivial, due to the dependency of the registered signal to various factors that may vary - in space or in time - during a given experiment, such as the excitation intensity, the probe concentration and the photo-bleaching of the fluorophores (Becker (2012)). Moreover, fluorescence intensity approaches are unable to distinguish different fluorescent probes with very similar fluorescence spectra, and different fractions of a fluorophore in different states of interaction with its molecular environment.
Notably, to address these limitations, it is possible to explore an additional characteristic of the fluorescence emission process: the excited-state lifetime or more precisely the temporal fluorescence decay function. Indeed, when a fluorophore is excited thanks to the absorption of a photon, it can return to the ground state following different pathways: by fluorescence (rate r_{fl}), by internally converting the absorbed energy of the photon into heat (rate r_{eh}), by passing the energy to its molecular environment (rate r_{ee}), or by crossing into the triplet state and return to the ground state from there (rate e_{ts}). The resulting fluorescence decay may be described by means of a simple two-state model, accounting for the excited state (population

N_1) and for the relaxed state of the molecules (population N_0). For a homogeneous population of molecules, we can write the variation of the population N_1 as follows:

$$\frac{dN_1(t)}{dt} = -(r_{fl} + r_{eh} + r_{ee} + r_{ts})N_1(t). \tag{4.1}$$

The solution is a single exponential function, which time constant τ is the reciprocal sum of the rate constants of all possible return paths, and it is called the fluorescence lifetime:

$$N_1(t) = e^{-(r_{fl}+r_{eh}+r_{ee}+r_{ts})t} = e^{-t/\tau}. \tag{4.2}$$

Conversely to the fluorescence intensity, the fluorescence lifetime is not sensitive to the probe concentration; indeed it depends mostly on the type of the molecule, its conformation and on its interactions with the molecular environment (Becker (2012)). In other words, the fluorescence decay function - among all the fluorescence parameters - provides the most direct insight into the molecular interactions of a fluorophore with its biological environment. Moreover, it is less susceptible to artefacts arising from scattered light, photobleaching, non-uniform illumination of the sample, or excitation intensity variations (Suhling et al. (2015)). The fluorescence lifetime information can be exploited, in the context of laser scanning microscopy architectures, to generate an image in which each pixel represents the average lifetime of the fluorescence registered in a given position of the raster scan. Since this approach leverages the fluorescence lifetime and not the fluorescence intensity, it is named fluorescence lifetime imaging (FLIM).

From a technical perspective, FLIM microscopes aim to measure the fluorescence intensity signal - obtained in the framework of a pulsed excitation scheme - with a temporal resolution in the order of picoseconds. It is possible to achieve this goal *via* two different approaches, commonly known as the temporal domain approach and the digital frequency domain (DFD) approach. Very briefly, the temporal domain approach relies on time-correlated single photon counting (TCSPC) hardware to measure the delay of a given photon-detection event with respect to the related excitation pulse; the DFD approach additionally exploits the concept of the heterodyne waves to great benefit (Gratton et al. (1984)), relaxing the technological requirements to obtain a target temporal resolution. Regardless of the method of choice, the raw data collected during a FLIM measure can be represented as a three-dimensional array which contains the histogram of the delay of the detected photons with respect to the excitation event, for every pixel of the two-dimensional scan. Each temporal histogram represents the average fluorescence decay function of all the fluorophores excited within the considered pixel. Notably, by fitting the histogram with the appropriate mono- or multi-exponential

model, it is possible to extract the single or multiple lifetime values that describe the temporal behaviour of the observed fluorophore populations. Ultimately, such lifetime values are used to populate the fluorescence lifetime image.
A major challenge for the successful fit of the time histogram is the need to acquire sufficient photons to achieve the precision in lifetime determination required to describe the temporal behaviours of the observed populations (Warren et al. (2013)). As an example, it has been calculated that tens of thousands of photons are required to accurately fit a bi-exponential decay (Köllner and Wolfrum (1992)); this is at least one order of magnitude greater the number available for each pixel from typical live cell experiments. Since the rate at which fluorescence photons can be practically detected is usually limited to avoid photobleaching, the minimum detected photon number requirements for FLIM can increase the image acquisition time, eventually neglecting high content/throughput assays. Therefore, any strategy to increase the number of registered photons - avoiding the increase of the excitation laser intensity - would be extremely beneficial in the context of fluorescence lifetime imaging.
To partially satisfy this need, in the next Section I will introduce fluorescence lifetime image scanning microscopy (FLISM), i.e., the combination of our ISM platform with the fluorescence lifetime assay, and I will show through experiments - on fixed and live cells - its superior performances with respect to traditional confocal FLIM analysis.

4.2 Fluorescence lifetime image scanning microscopy

In the previous Chapter we presented our platform for imaging scanning microscopy based on the SPAD array detector. As described, our implementation upgrades the traditional confocal setup by substituting the single element detector with the SPAD array: this simple modification in the detection arm leads to ISM results with improved spatial resolution and signal to noise ratio, with respect to the confocal counterparts. Here, we decline the same benefits within the framework of fluorescence lifetime imaging, and we call the proposed architecture fluorescence lifetime image scanning microscopy (FLISM). We obtain the FLISM results leveraging the fast temporal impulse response function (low photon-timing jitter) of each element of the SPAD array detector, coupled with a multi-channel TCSPC card; such results show:

- increased spatial resolution with respect of the traditional FLIM counterpart, allowing to better discriminate the observed structures;

- increased signal to noise ratio. Whereas the SNR enhancement provided fluorescence intensity images with better contrast, in the context of lifetime imaging it assumes a more prominent role. Indeed, as described in the previous Section, the amount of detected photons is a critical aspect of any FLIM measurement, since it directly impacts the precision in the lifetime determination. Hence, in the context of lifetime imaging, the SNR enhancement provided by the FLISM approach translates in images with increased lifetime accuracy, when compared to the conventional confocal FLIM counterpart.

In other words, although the overall photon timing jitter is the same for FLIM and FLISM (i.e., same laser, same detector, same optics, and same multi-channel time-to-digital converter), the precision of FLISM is higher due to the higher SNR.

4.3 Material and methods

4.3.1 Setup

The setup used to perform the measurements presented in this Chapter is similar to the one described in Chapter 3. Additionally, to perform the lifetime measurements, we connected the five central elements of the SPAD array module to a multi-channel time-to-digital converter (TDC) with a temporal jittering of 80 ps (FWHM) and a maximum sustained data rate of 8.5 million tags per second (Time Tagger 20; Swabian Instruments), working in the so-called time-tag modality. We appropriately delayed the electronic trigger output signal provided by the driver of the excitation laser (485 nm or 640 nm, depending on the imaging experiment) running at 80 MHz *via* a picosecond electronic delayer (Picosecond Delayer; Micro Photon Devices) and used the output as the reference signal (sync) for the time-resolved measurement. The *carma* microscope control architecture generated the pixel, line and frame clocks, which we sent to the multi-channel TDC. We used a custom C# software module to read the stream of events outputted by the Time Tagger 20 card, each of them labeled with the corresponding inputs (sync, pixel, line, frame or element 1–5) and the time of arrival (time tag). To reduce the data rate, the card discards all synchronization events (80 MHz) except those that follow after one of the photon-detection low-rate events, thereby forming a reduced time-tag stream. Once the time-tag stream has been created, the micro-time (delay from the sync signal) of each photon event is calculated, and the scanned lifetime histogram images (128 bins, 100 ps each), one for each SPAD element, are generated. The result is a series of three-dimensional scanned images in which the third dimension represents the lag time (photon-arrival time) of

the histogram. Relative delays between the different elements of the detector are corrected through measurement of their impulse-response functions.

4.3.2 Adaptive pixel reassignment in time-resolved measurements

The usage of the SPAD array detector in the laser scanning microscope results in an increase of the dimensionality of the collected data: performing a scan of the sample with the ISM implementation described in Chapter 3 results in a set a 4D data set, indeed together with the conventional spatial scanning dimension, we have two more spatial dimensions which denote the position in the detector array, in our case 5×5. In the context of traditional fluorescence lifetime imaging, a single scan results in a 3D array in which the third dimension represents the temporal histogram of the detected photons with respect to the excitation event. When substituting the single detector with the SPAD array detector, the single scan provide a 5D data set, where the fifth dimension is the temporal histogram. For a given 5D time-resolved dataset, we obtained the final FLISM image using the following protocol:

1. for each element of the detector array (x',y') the 3D image (x,y,t) is integrated along the temporal dimension t, to take into consideration all detected photons, thus neglecting the temporal information of the time of arrival;

2. we apply the phase correlation algorithm described in Chapter 3 on this set of 4D images, to derive - directly from the data - all the shift values; the higher photon counts related to the time-integrated images result in more reliable shift values;

3. lastly, we consider the original pool of 5D images and we perform - for each temporal bin of the histogram - the adaptive pixel reassignment operation described in Chapter 3 using the shift values obtained in the step 2.

This protocol leads to the raw FLISM 3D array. Notably, by summing the photons for each temporal bin - instead of reassigning them - we obtain the raw FLIM counterpart. We finally analysed the 3D FLISM/FLIM dataset in the time domain with the FLIMfit software tool developed at Imperial College London (https://www. flimfit.org). In short, the software fits the fluorescence decay trace (i.e., the photon-arrival-time histogram) of each pixel with a single or multi-exponential function to extract the fluorescence lifetime τ. Alternatively, we analysed the 3D FLISM/FLIM data in the frequency domain by using a custom software able to transform the decay trace of each pixel into a pair of values, the phasor point (g,s), for display on the polar/phasor plot (see Chapter 6). The phasor values are the normalized

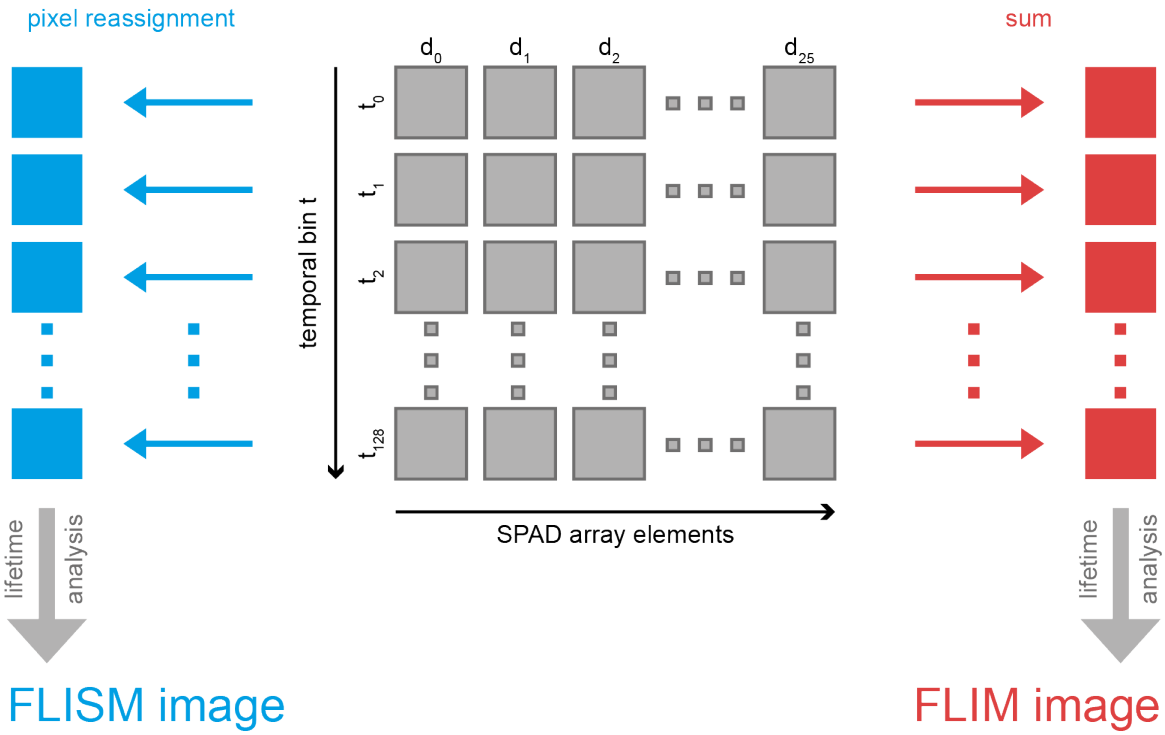

Figure 4.1 **Adaptive pixel reassignment in time-resolved measurements**. Schematic of the protocol used to compare the the traditional FLIM raw image and our FLISM raw image. FLIM (right): the photons for each temporal bin are summed. FLISM (left): the photons for each temporal bin are reassigned using our adaptive pixel reassignment algorithm. To assess the shift values, we apply the phase correlation algorithm on the images obtained integrating along the temporal dimension.

cosine (g) and sine (s) transforms at modulation frequencies equal to the repetition rate of the excitation laser.

4.4 Results

We tested the proposed fluorescence lifetime image scanning microscopy architecture first by imaging tubulin filaments, labeled with Abberior STAR red, in a fixed HeLa cell (Fig. 4.2). The results underline the increased spatial resolution of the FLISM result when compared to the traditional FLIM counterpart, obtained in the "open pinhole" configuration. We then analysed both the raw FLIM and the raw FLISM images by attempting to fit the fluorescence decay for each pixel with a single exponential function: notably, the fitting procedure of the raw FLIM data related to the "close pinhole" configuration fails for various pixels, due to the low signal to noise ratio. As expected, the fitting success rate for the raw FLISM image is

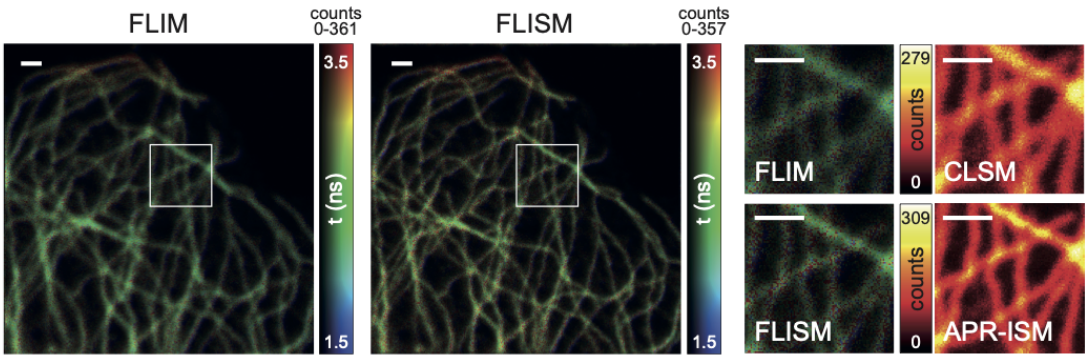

Figure 4.2 **Fluorescence lifetime image scanning microscopy on fixed cell** Side-by-side comparison of the fitting-based FLIM (left, 1.4 A.U.) and the FLISM (right) images of tubulin labeled with Abberior STAR red, and magnified views of the white boxes paired with the respective intensity images. Pixel-dwell time: 100 μs. Pixel-size: 30 nm. Image format: 500 $\times$ 500 pixels. Excitation power P_{exc} = 500 nW.

significantly increased.

Next, we imaged live cells with an ANEP polarity-sensitive membrane probe, which changes in the fluorescence lifetime allows to distinguish domains of the membrane in liquid-ordered and liquid-disordered phases. Also in this case, the increased signal to noise ratio granted by the FLISM analysis grants an higher success ratio of the fitting procedure (Fig. 4.3a). Moreover, we analysed the same FLISM dataset *via* the phasor approach. Again, thanks to the higher SNR, the points in the FLISM phasor plot are located more precisely than those in the FLIM counterpart, which clearly allowed for better separation of domains in liquid-ordered and liquid-disordered phases (Fig. 4.3b).

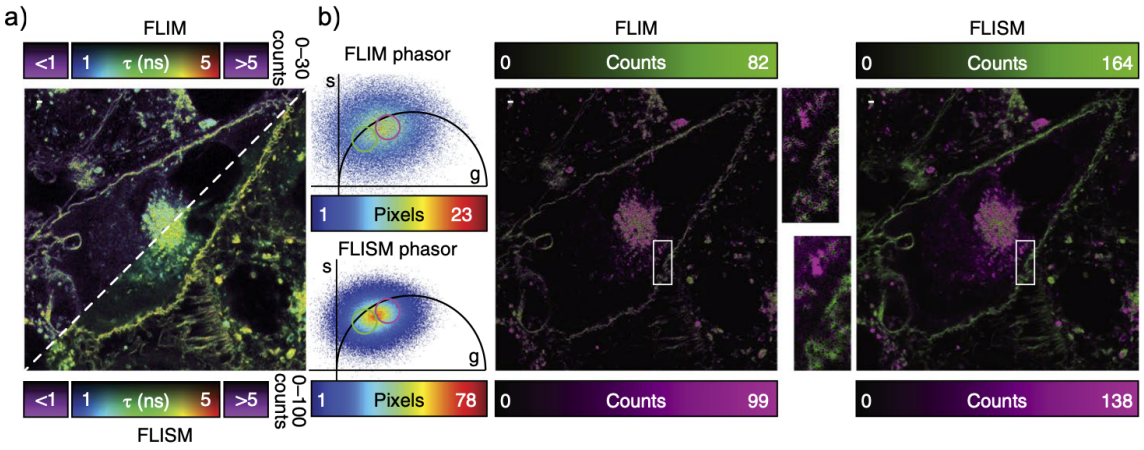

Figure 4.3 **Fluorescence lifetime image scanning microscopy for live cell imaging** (a) Comparison of the fitting-based FLIM (top-left, 0.25 A.U.) and the FLISM (bottom-right) images of membrane labeled with ANEP. Pixels in purple correspond to out-of-range FL values, which likely results from artefacts during the exponential fitting analysis. All pixel are analysed. (b) Comparison of the phasor plots for the FLIM and FLISM datasets (only the points/pixels above the background level are represented, i.e. 5 and 15 photons, respectively for the FLIM and FLISM images), together with the images obtained by back-projecting the points within the purple (long lifetime) and red (short lifetime) circles and magnified views of the white boxes ($n = 10$ experiments).